

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2004, 00:50:22 ; Search time 3791 Seconds
(without alignments)
10586.856 Million cell updates/sec

Title: US-10-042-894A-7

Perfect score: 1344

Sequence: 1 gcacaggtcagtcctgcac.....ataaaaaaaaaaaaaaaa 1344

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

1: em_estba:*
2: em_esthum:*
3: em_estim:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_estl:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vri:*
28: gb_gss1:*
29: gb_gss2:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	921	68.5	960	28	CC344833
2	839.4	62.5	923	29	CG240397
3	762.2	56.7	818	29	CG453578
4	752	56.0	873	28	CC344824

5	644.8	48.0	867	29	CC724960
6	620	46.1	801	13	CA134480
7	613.2	45.6	719	29	CG284572
8	589.4	43.0	591	29	CG776236
9	564.6	42.0	656	14	CA245555
10	559	41.6	722	13	CA130685
11	557.8	41.5	3374	11	AY109355
12	555	41.3	671	14	CA202412
13	540.8	40.2	722	29	CG724950
14	535.6	39.9	593	28	BZ774817
15	525.6	39.1	744	13	CA130686
16	520.8	38.0	960	29	CG073495
17	500	37.2	907	29	CG240387
18	485	36.1	955	28	CC384766
19	440.4	32.8	1191	14	CK211413
20	435.6	32.4	794	10	BF617713
21	426.2	31.7	575	29	CG734025
22	423	31.5	551	28	BH408057
23	420.4	31.3	619	14	CA204213
24	376.2	28.0	842	28	BZ736665
25	376.2	28.0	866	29	CG367966
26	376.2	28.0	939	29	CG367979
27	372.8	27.7	588	13	BU999708
28	351	26.1	927	29	CG284580
29	339.2	25.2	876	29	CG073494
30	334.4	24.9	431	29	CG776505
31	334.2	24.9	557	10	BF277833
32	333.6	24.8	448	28	CC028437
33	332.8	24.8	1099	14	CK215908
34	324.8	24.2	1123	14	CK214494
35	319	23.7	453	12	BI956405
36	311.2	23.2	433	12	BM099879
37	310.4	23.1	317	29	CG194403
38	309.8	23.1	598	29	CG386200
39	298	22.2	875	29	CC673013
40	297.2	22.1	620	13	BQ820191
41	233.6	21.8	409	12	BM099878
42	288.8	21.5	478	12	BG417083
43	277	20.6	306	28	CC023125
44	267.2	19.9	606	13	BU051170
45	250.4	18.6	506	29	CG366904

ALIGNMENTS

RESULT 1
CC344833
LOCUS CC344833 960 bp DNA linear GSS 16-MAY-2003
DEFINITION OGIAQ20TV_ZM_0.7_1.5_KB_Zea_mays_genomic_clone_ZMMMa0359C15,
genomic survey sequence.
ACCESSION CC344833
VERSION CC344833.1 GI:30814239
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 960)
AUTHORS Whitlaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nurnberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

SUMMARIES

```
FEATURES
  source
    1. 960
    /organism="Zea mays"
    /mol_type="genomic DNA"
    /strain="B73"
    /db_xref="taxon:4577"
    /clone="ZMBMa0359C15"
    /clone_lib="ZM 0.7 1.5 kb"
    /note="Vector: pBSK-; Site 1: HincII; 0.7-1.5 kb
    methylation filtered genomic DNA library"

ORIGIN
  Query Match      68.5%; Score 921; DB 28; Length 960;
  Best Local Similarity 100.0%; Pred. No. 1.4e-131;
  Matches 921; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 GTCCCTTCCCATACATGTCGAGCTCCAGCTCCACCGCGGAGACCAAGTCGCGGCGCAC 94
DB 1 GTCCCTTCCCATACATGTCGAGCTCCAGCTCCACCGCGGAGACCAAGTCGCGGCGCAC 60
QY 95 GCGCTCCGCGAGAGCTGGCGCGCTCATCGAGGCTCGGCGCTCTCTACAGCGCG 154
DB 61 GCGCTCCGCGAGAGCTGGCGCGCTCATCGAGGCTCGGCGCTCTCTACAGCGCG 120
QY 155 TCCAGCGCGGACCGTGGGAGGACAGAGTGCGCTTCTATGAGGCGTTCTCGGCCACG 214
DB 121 TCCAGCGCGGACCGTGGGAGGACAGAGTGCGCTTCTATGAGGCGTTCTCGGCCACG 180
QY 215 CGCGCTCCGCGCGCATCCAGAGACACTTCTTCCCGCGGTTTCCAGCGAGCGACTCC 274
DB 181 CGCGCTCCGCGCGCATCCAGAGACACTTCTTCCCGCGGTTTCCAGCGAGCGACTCC 240
QY 275 TCCCGACGAGCGAGCGCGGAGCGAGCTCCAGCTCTCGTCTCGAGGACCTCCCTCG 334
DB 241 TCCCGACGAGCGAGCGCGGAGCGAGCTCTCGTCTCGAGGACCTCCCTCG 300
QY 335 CGGGTTTCAGCGCGCTCGTCCGAGACATCAAGATCGCGCGCTCATCGTGGCCACCGA 394
DB 301 CGGGTTTCAGCGCGCTCGTCCGAGACATCAAGATCGCGCGCTCATCGTGGCCACCGA 360
QY 395 GTTCGCGGAGCGCTTACATCGCGAGTGTCTCGCAAGAGCCGCGGAGACAGAGGTTC 454
DB 361 GTTCGCGGAGCGCTTACATCGCGAGTGTCTCGCAAGAGCCGCGGAGACAGAGGTTC 420
QY 455 TGCTCGAGTTCCGCTTCCGCGCTCCGAGTGTCTCGCGCGCGAGCGCGGTTGCGCGA 514
DB 421 TGCTCGAGTTCCGCTTCCGCGCTCCGAGTGTCTCGCGCGCGAGCGCGGTTGCGCGA 480
QY 515 CGAGCGCGCGGAGTGAAGGCCATGGACACCGCGCGCTCCCGCGCTGCTCCGCGCGT 574
DB 481 CGAGCGCGCGGAGTGAAGGCCATGGACACCGCGCGCTCCCGCGCTGCTCCGCGCGT 540
QY 575 AGTGTATCTCCGTGCGCGAGGGGATGAGTGTGCGCTCCCGCGCGGTTGTCAGGAG 634
DB 541 AGTGTATCTCCGTGCGCGAGGGGATGAGTGTGCGCTCCCGCGCGGTTGTCAGGAG 600
QY 635 GAAAGGTGGAGTCTTGTCACAGCTGCGGAGCTCAAGCGCTGGTTCGAGGAGCAGATC 694
DB 601 GAAAGGTGGAGTCTTGTCACAGCTGCGGAGCTCAAGCGCTGGTTCGAGGAGCAGATC 660
QY 695 TGTTCCATCTTACTCGGCGTCAATCTTCTGGGCTATGATCTGCTGAGTTCGAGTCGAG 754
DB 661 TGTTCCATCTTACTCGGCGTCAATCTTCTGGGCTATGATCTGCTGAGTTCGAGTCGAG 720
QY 755 GCGGAGATGGGGTGGGTTGAGCTGAGCTGGTGAATTCGCCATGTGGCGGAGGTG 814
DB 721 GCGGAGATGGGGTGGGTTGAGCTGAGCTGGTGAATTCGCCATGTGGCGGAGGTG 780
QY 815 ATGGGTTGATTCACACACTTCTCGGCGGCTCTGCTCGTGTATCAAGTTCGTTTCTG 874
DB 781 ATGGGTTGATTCACACACTTCTCGGCGGCTCTGCTCGTGTATCAAGTTCGTTTCTG 840
QY 875 ACATTGTTCCGAGACTCTCTCAATAGCGAGCTTTGGGTCTCTTCTTAAGAGAGGATCTGG 934

Db 841 ACATTGTTCCGAGACTCTCTCAATAGCGAGCTTTGGGTCTCTTAAAGAGAGATCTCTGG 900
QY 935 CATTTTCGATTTGATAACAAAG 955
Db 901 CATTTTCGATTTGATAACAAAG 921

RESULT 2
CG240397 923 bp DNA linear GSS 22-AUG-2003
LOCUS OGYC127TV ZM 0.7 1.5 KB Zea mays genomic clone ZMBMa0789E05,
DEFINITION genomic survey sequence.
ACCESSION CG240397
VERSION CG240397.1 GI:34140283
SOURCE GSS.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 923)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
CONSORTIUM Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other_GSSs: OGYC127TH
Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@cigr.org
Seq primer: TP
Class: sheared ends.
FEATURES
  source
    1. 923
    /organism="Zea mays"
    /mol_type="genomic DNA"
    /strain="B73"
    /db_xref="taxon:4577"
    /clone="ZMBMa0789E05"
    /clone_lib="ZM 0.7 1.5 KB"
    /note="Vector: pBSK-; Site 1: HincII; 0.7-1.5 kb
    methylation filtered genomic DNA library"

ORIGIN
  Query Match      62.5%; Score 839.4; DB 29; Length 923;
  Best Local Similarity 97.3%; Pred. No. 4.5e-119;
  Matches 875; Conservative 0; Mismatches 21; Indels 3; Gaps 2;

QY 42 TCCCATACATATGTCGAGACTTCCACCGCGGAGACCAAGTCGCGGCGCTCCAGGC 101
DB 24 TCCCATACATATGTCGAGACTTCCACCGCGGAGACCAAGTCGCGGCGCTCCAGGC 83
QY 102 CGCAGCAAGTCGCGCGCTCATCGAGGCTCCGCGCTCTTCTACAGCGCTCCAGGC 161
DB 84 CGCAGCAAGTCGCGCGCTCATCGAGGCTCCGCGCTCTTCTACAGCGCTCCAGGC 143
QY 162 CGCGACCGTGGGAGCAGAGGTCGCTTCTATGAGCGCTTCTCCGCCACCGCGCGT 221
DB 144 CGCGACCGTGGGAGCAGAGGTCGCTTCTATGAGCGCTTCTCCGCCACCGCGCGT 203
QY 222 CCGCGCGCGATCCGAGACACTTCTTCCCGCGGTTTCCAGCGAGGACTCTCTCCCGAC 281
DB 204 CCGCGCGCGATCCGAGACACTTCTTCCCGCGGTTTCCAGCGAGGACTCTCTCCCGAC 263
QY 282 CGAGCGGAGCGCGGAGCGGATCTCTCACTCGTCTCTCGAGGCTCTCTCGGCGGT 341
DB 264 CGAGCGGAGCGCGGAGCGGATCTCTCACTCGTCTCTCGAGGCTCTCTCGGCGGT 323
QY 342 TCAGCGCGCTCGCTCGCAGACATCAAGATCGCGGCGCTATCAGTGGCGCACCGAGTTGCC 401
```



```

ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
source
1..873
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZM0359C15"
/clone="ZM0359C15"
/notes="Vector: pBCSK-; Site 1: HindII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 56.0%; Score 752; DB 28; Length 873;
Best Local Similarity 100.0%; Pred. No. 1.1e-105; Indels 0; Gaps 0;
Matches 752; Conservative 0; Mismatches 0;
QY 204 CTCGCCACGCGCGCGTCCGCGCGCATCCGAGACACCTTCTCCCGCGTTCCACGG 263
DB 873 CTCGCCACGCGCGCGTCCGCGCGCATCCGAGACACCTTCTCCCGCGTTCCACGG 814
QY 264 CAGCGACTCTCCACAGCGCGAGCGCGCGGAGCGCGATCTCTCTCTCTCTCTCTCT 323
DB 813 CAGCGACTCTCCACAGCGCGAGCGCGCGGAGCGCGATCTCTCTCTCTCTCTCTCT 754
QY 324 CGACCTCTCTCGCGGGTTTCAGCGCGCGTTCGCGCGAGCATCAAGATCGCGCGCATAC 383
DB 753 CGACCTCTCTCGCGGGTTTCAGCGCGCGTTCGCGCGAGCATCAAGATCGCGCGCATAC 594
QY 384 GTGGCCACCGAGTTCGCGGAGCCCTACATCGCCAAAGTGCCTCGCCAAAGACCGCGGAC 443
DB 693 GTGGCCACCGAGTTCGCGGAGCCCTACATCGCCAAAGTGCCTCGCCAAAGACCGCGGAC 634
QY 444 CAGGAGCTTCTCTCGGATTCGCGTCTCGCGCTCGAGTCTCGCGCGCGCGCGCGCGCG 503
DB 633 CAGGAGCTTCTCTCGGATTCGCGTCTCGCGCTCGAGTCTCGCGCGCGCGCGCGCGCG 574
QY 504 CTTGTGCGCGACGCGCGCGCGGATGAGGCGATGAGCGCGCGCGCGCGCGCGCGCGCG 563
DB 573 CTTGTGCGCGACGCGCGCGCGGATGAGGCGATGAGCGCGCGCGCGCGCGCGCGCGCG 514
QY 564 GTCGCGCGCTACGTGTATTCGTTTCCGACGAGGCGGATGAGTGTGCGTCTCGCGCGCG 623
DB 513 GTCGCGCGCTACGTGTATTCGTTTCCGACGAGGCGGATGAGTGTGCGTCTCGCGCGCG 454
QY 624 GGTGTACGAGGAAAGGTGGAGTCTTGTACAGCTCGCGGAGTCAAGCGGTGGTTCGA 583
DB 453 GGTGTACGAGGAAAGGTGGAGTCTTGTACAGCTCGCGGAGTCAAGCGGTGGTTCGA 394
QY 684 GGAGCAGACTCTCTTCCACTTCTACTCGCGGTTCGATTTCTTCTGGGTATGATGCTGCTGC 743
DB 393 GGAGCAGACTCTCTTCTTCTTCTACTTCTCGGCTCGATTTCTTCTGGGTATGATGCTGCTGC 334
QY 744 AGTCGACGAGCGGAGATGGGGTGGGTTGACGGTGAAGCTGGTGGACTTTGCCCATGT 803
DB 333 AGTCGACGAGCGGAGATGGGGTGGGTTGACGGTGAAGCTGGTGGACTTTGCCCATGT 274

```

```

QY 804 GCGCCAGGGTGATGGGTGATTGACACAACTTCTCTGGCGGGCTCTGCTCGTGATCAA 863
DB 273 GCGCCAGGGTGATGGGTGATTGACACAACTTCTCTGGCGGGCTCTGCTCGTGATCAA 214
QY 864 GTTCGTTTCTGACATGTTCCGGAGACTCCTCATACGAGCGCTTTGGGTCTTCTTAAGA 923
DB 213 GTTCGTTTCTGACATGTTCCGGAGACTCCTCATACGAGCGCTTTGGGTCTTCTTAAGA 154
QY 924 GAGGATCTCTGGCATTTTCGATTTGATAACAAAG 955
DB 153 GAGGATCTCTGGCATTTTCGATTTGATAACAAAG 122

RESULT 5
CC724960 867 bp DNA linear GSS 23-JUN-2003
OGLAB23TV ZM 0.7 1.5 KB Zea mays genomic clone ZMVBMA0307C22,
DEFINITION
genomic clone survey sequence.
ACCESSION
CC724960 GI:32143893
VERSION
CC724960.1
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGLAB23TH
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
source
1..867
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMVBMA0307C22"
/clone_lib="ZM 0.7 1.5 KB"
/notes="Vector: pBCSK-; Site 1: HindII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 48.0%; Score 644.8; DB 29; Length 867;
Best Local Similarity 89.4%; Pred. No. 3.1e-89;
Matches 753; Conservative 0; Mismatches 27; Indels 62; Gaps 3;
QY 2 CACGAGGTGAGTCCGTCACCCCTCGCGCCATAGTCCCTTCCCCATACCATGTCGACC 61
DB 87 CACCATGCGCACCCGTCACCCCTTCTGCTCCATAG-----TCCCCATACCATGTCGACC 140
QY 62 TCACCCCGCGGAGCACCAAGTCCGCGGCCACCGCGCTCCCGCAGCAGCTGGGCCCGC 121
DB 141 TCACCCCGCGGAGCACCAAGTCCGCGGCCCTCCCGCGCTCCCGCAGCAGCTGGGCCCGC 200
QY 122 TCATCGAGCGTCCGCGCTCTTCTAAGCGCGCTCCAGCGCGCGCGCGCGCGCGCGCAG 181
DB 201 TCATCGAGCGTCTGGCGCTCTTCTAAGCGCGCTCCAGCGCGCGCGCGCGCGCGCGCAG 260
QY 182 AGGTGCGCTTCTATGAGCGGTTCTCGCGCCACCGCGCGCTCCCGCGCGCGCGCATCCGAGACA 241
DB 261 AGGTGCGCTTCTATGAGCGGTTCTCGCGCCACCGCGCGCTCCCGCGCGCGCGCATCCGAGACA 320

```



```
Db      676 TGGTCCACCTTTACTCGGGGGCGATTCTTCTGGGCTATGAAGCTAGTGCCGTTGCAACA 735
QY      755 GCGG-AGATGGGGTGGGGTGACGATGAGCTGGGTGGACTTTGCCCATGTGGCCGAGGGT 813
Db      736 CCGGAAGGTAAAGGTGGGGTGAAGGTAAACATGGGGGCCCTTTGCCATGTGGCCCAAGG 795
QY      814 GATGGG 819
Db      796 GAAGGG 801

RESULT 7
LOCUS   CG284572
DEFINITION CG284572 719 bp DNA linear GSS 25-AUG-2003
GENOMIC SURVEY SEQUENCE.
ACCESSION CG284572
VERSION   CG284572.1 GI:34198786
KEYWORDS GSS.
SOURCE   Zea mays
ORGANISM Zea mays
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
          clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 719)
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
          Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
          Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
          Consortium for Maize Genomics
          Unpublished (2002)
          Other GSSs: OGMIF61TV
          Contact: Cathy Whitelaw
          TIGR
          9712 Medical Center Drive, Rockville, MD 20850, USA
          Tel: 301-838-5843
          Fax: 301-838-0208
          Email: whitelaw@tigr.org
          Seq primer: TR
          Class: sheared ends.

FEATURES             source
    source
    1..719
    /organism="Zea mays"
    /mol_type="genomic DNA"
    /strain="B73"
    /db_xref="taxon:4577"
    /clone="ZMBM0584K02"
    /clone_lib="ZM_0.7_1.5_KB"
    /note="Vector: pBCKS-; Site_1: HincII; 0.7-1.5 kb
    methylation filtered genomic DNA library"

ORIGIN
Query Match          45.6%; Score 613.2; DB 29; Length 719;
Best Local Similarity 90.4%; Pred. No. 2.3e-84;
Matches 700; Conservative 0; Mismatches 18; Indels 56; Gaps 2;

QY      50 CCATGTCGACCTCCACCGCGGAGCACCAAGTCGCGCGCCACCGCGCTCCGCCAGCA 109
Db      1 CCATGCCGACCTCCACCGCGCGGAGCACCAAGTCGCGCGCTCCGCCAGCA 60
QY      110 AGCTGGGCGCGCTCATCGACGCTCGGCTCTTCTACAGCGCTTCAGGCGCGCGGACC 169
Db      61 AGCTGGGCGGACCTCATCGACGCTCTGGGCTCTTCTACAGCGCTTCAGGCGCGGACC 120
QY      170 GTGGGAGACAGGCTGCGCTTCTATGAGCGCTTCTCGGCCACGCGCGCTCCCGGCC 229
Db      121 GTGGGAGACAGGCTGCGCTTCTATGAGCGCTTCTCGGCCACGCGCGCTCCCGGCC 180
QY      230 GCATCCGAGACACCTTCTTCCCGCGTTCACCGGACGCGACTCTCTCCCAACCGAGGCGC 289
Db      181 GCATCCGAGACACCTTCTTCCCGCGTTCACCGGACGCGACTCTCTCCCAACCGAGGCGC 240
QY      290 AGCCGGGAGCGGCTCTCTACCTCGTCTCTGACGACCTCTCGGGGGTTTCAGGCGC 349
```

```
Db      241 AGCCGGGAGCGCATCCGACCTCGTCTCGACGACCTCTCTCGGGGTTTGAGGCGC 300
QY      350 CTGGTTCGAGACATCAAGATCGCGGCATCACTGTCGCCACCGAGTTCGCGGAGCCCT 409
Db      301 CTGGTTCGAGACATCAAGATCGTGCATCAGT----- 336
QY      410 ACATCGCCCAAGTGCCTCGCCCAAGGACCGCGGACCCACGAGCGTTCTGTCGGATTCGCG 469
Db      337 -----GACCACGAGCGTCTGTCGGATTCGCG 365
QY      470 TCTCGGCGTCCGAGTCTCGGCCCCGAGGCGCGCTGTGGCGGACGAGCGCCCGGAGG 529
Db      366 TCTCGGCGTCCGAGTCTCGGCCCCGAGGCGCGCTGTGGCGGACGAGCGCCCGGAGG 425
QY      530 TGAAGGCATGAGACACCGCGCGCTCGCGCGCTCGCGCGCTACCTGTGTCATCCG-TT 588
Db      426 TGAAGGCTATGACATTGTCGCGCGTCCGCGCGCTCGCGCGCTACCTGTGTCATCCGTT 485
QY      589 GCCGACGAGGCGATGGAAGTGTGCGCTCGCGCGCGGCTGTACGGAGGAAAAGGTGAGTGC 648
Db      486 GCCGACGAGGCGATGGAAGTGTGCGCTCGCGCGCGGCTGTACGGAGGAAAAGGTGAGTGC 545
QY      649 TTGTCACAGTTCGCGGAGCTCAAGCGGTGTTGAGGAGGAGACTCTGTTCCACTTCTAC 708
Db      546 TTGTCACAGTTCGCGGAGCTCAAGCGGTGTTGAGGAGGAGACTCTGTTCCACTTCTAC 605
QY      709 TCGGCGTCTGATCTTCTGGGCTATGATGCTGCTCAGTCGCGAGCGGCGGAGATGGGGT 768
Db      606 TCGGCGTCTGATCTTCTGGGCTATGATGCTGCTCAGTCGCGAGCGGCGGAGATGGGGT 665
QY      769 GGGGTGACGCTGAAGCTGTGAGACTTTTCCCATGTGGCGCGGGGTGATGGGGTG 822
Db      666 GGGGTAAACAGTGAAGTGTGAGACTTTTCCCATGTGGCGCGGGGTGATGGGGTG 719

RESULT 8
LOCUS   CG776236
DEFINITION 1123005E08.x1 1123 - RescueMu Grid L Zea mays genomic, genomic survey sequence.
ACCESSION CG776236
VERSION   CG776236.1 GI:38034108
KEYWORDS GSS.
SOURCE   Zea mays
ORGANISM Zea mays
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
          clade; Panicoideae; Andropogoneae; Zea.
          1 (bases 1 to 591)
          Walbot,V.
          Maize genomic sequences found using engineered RescueMu transposon
          Unpublished (2001)
          Contact: Walbot V
          Department of Biological Sciences
          Stanford University
          855 California Ave, Palo Alto, CA 94304, USA
          Tel: 650 723 2227
          Fax: 650 725 8221
          Email: walbot@stanford.edu
          Plate: 1123005 column: 3
          Class: transposon-tagged.
          Location/Qualifiers
            1..591
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /cultivar="mixed background W23/A188/B73/X55"
            /db_xref="taxon:4577"
            /dev_stage="leaf"
            /tissue_type="adult"
            /lab_host="DH10B"
            /clone_lib="1123 - RescueMu Grid L"
            /note="Organ: leaf; Vector: RescueMu (engineered from
            pBlueScript backbone); Site_1: BamHI; Site_2: BglIII;
```


QY 577 GTGTCATCCGTGCGGACGAGGGAGTGAAGTGTGCGCTCGCGCGCGCGGTGTACGGAGG- 635
 Db 558 GTTTCATCCGTGCGGACGAGGGAGTGAAGTGTGCGCTCGCGCGCGCGGTGTACGGCGGC 617

QY 636 AAAAGGTGGAGTCTGTGTCACAGCTCGCGAGTCAAGGC 674

Db 618 AAAAGGGGAGTCTGTGTCACAGCTCGCGAGCTTAAGGC 656

RESULT 10
 LOCUS CAL130685 722 bp mRNA linear EST 24-SEP-2003
 DEFINITION SCCRR11004G05.g RT1 Saccharum officinarum cDNA clone SCCRR11004G05
 5', mRNA sequence.

ACCESSION CAL130685

VERSION CAL130685.1 GI:35013914

KEYWORDS EST.

SOURCE Saccharum officinarum

ORGANISM Saccharum officinarum

REFERENCE Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

AUTHORS clade; Panicoideae; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

TITLE 1 (bases 1 to 722)

JOURNAL The libraries that made SUCST

COMMENT Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

Contact: Arruda P

Centro de Biologia Molecular e Engenharia Genetica

Universidade Estadual de Campinas

Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137

Fax: 55 19 3788 1089

Email: parruda@unicamp.br

Clone distribution: clone distribution information can be found

through the Brazilian Clone Collection Center (BCCC) at

http://www.bcccenter.fcav.unesp.br

Plate: 004 row: G column: 05

Seq primer: T7 Promoter Primer.

Location/Qualifiers

1. .722

/organism="Saccharum officinarum"

/mol_type="mRNA"

/db_xref="taxon:4547"

/clone="SCCRR11004G05"

/lab_host="DH10B"

/clone_lib="RTL"

/note="Organ: Root tips (0.3cm-long) from adult plants;

Vector: pSport1; Site 1: SalI; Site 2: NotI; An

unidirectional cDNA library generated from [Root tips

(0.3cm-long) from adult plants]. cDNA was prepared from

polyA+ mRNA using SuperScript Plasmid System Kit

(Invitrogen). The double-strand cDNAs were fractionated

QY 37 CCCCTTCCCATACCATGTCGAGCTTCCACCGCGGAGCACCAGTCCGCGGCACCGC 96
 Db 16 CTTCTCGCCCTCACCATGTCGAGCTTCCACCGCGGAGCACCAGTCCGCGGCACCGC 74

QY 97 GCTCGGCGAGAGTGGCGGCGCTCATCGAGCGCTCGGCGCTTCTTACAGCGGCTC 156

Db 75 GCTTCGGCGCAACAGCTGGGTTCGCTCATCGAGCGCTTCTTACAGCGGCTC 134

QY 157 CAGCGCGCGGACCGTGGGAGCAGCAGGTCGCGCTTCTATAGGCGTTCCTCCGCCACCGC 216

Db 135 CAGTCCGGGACCCCGGGAGCAGAGCTCGCTTCTACGAGCGTTCTCCACCCACGCC 194

QY 217 GCCCTCCCGGCGGCATCCGAGACACCTTCTTCCCGGGTTCCACGGCAGCGACTCTC 276

Db 195 GCCGTCCCGGCGGCATCCGAGACACCTTCTTCCCGGGTTCCACGGCAGCGACTCTC 254

QY 277 CCACCGAGCGGAGCCCGGGAGCGGCATCCTCCTCGCTTCGAGGACTCTCTCGCG 336

Db 255 CCCACCGAGCGCGCGGAGCGGCATCCTCGGCATCGCTTCCTCGAGGACTCTCTCGCG 314

QY 337 GGGTTTCAGGGCGCCCTCGTCGAGACATCAAGATCGGCGGCATCACTGTCGCCACCGAGT 396

Db 315 GGGTTCGAGGGCGCCCTCGTCGCGCATCAAGATCGGCGGCATCACTGTCGCCACCGAGC 374

QY 397 TCGCGGAGCGCCCTACATCGCCAAAGTCTCGCCAAAGACCGCGGAGCACCAGGCTTCTG 456

Db 375 TCGCGGAGCGCCCTACGTCGCGCAAGTCTCGCCAAAGACCGCGGAGCACCAGGCTTCTG 434

QY 457 CTCGATTCGCGCTCCGGGTCGCGAGTCTCGGCGCCCGGAGCGCGCTGTCGGGCGCTAC 516

Db 435 CTCGATTCGCGCTCCGGGTCGCGAGTCTCGGCGCCCGGAGCGCGCTGTCGGGCGCTAC 494

QY 517 GAGCGCGCGGAGGTGAAGGCGCATGACACCGCGCGCTCGCGCGCTGCTCCGGGCGCTAC 576

Db 495 GAGCGCGCGGAGGTGAAGGCGCATGACACCGCGCGCTCGCGCGCTGCTCCGGGCGCTAC 554

QY 577 GTGTCATCCGTGCGGAGCGGAGTGAAGTCTCGCTTCGCGCGCGGCTGTA-CGGAGG 635

Db 555 CGTTTCATCCGTGCGGAGCGGAGTGAAGTCTCGCTTCGCGCGCGGCTGTA-CGGAGG 614

QY 636 AAAAGGTGGAGTCTGTGTCACAGCTCGCGAGCTCAAGCGCTGTCGAGGAGCAGTCT 695

Db 615 CAAAAGGGGAGTCTGTGTCACAGCTCGCGAGCTTAAAGCGTGGTGGAGGA-CAAACTCT 673

QY 696 GTTCCACTTCTACTCGCGGTCGATTCCTTCGGGCTATGATGCT 738

Db 674 GGATCACTT-TACTCGGGCGGATTCCTTCGGGCTATGATGCT 715

RESULT 11

AY109355

LOCUS Zea mays CL49_2 mRNA sequence. 3374 bp mRNA linear HTC 17-OCT-2002

DEFINITION Zea mays CL49_2 mRNA sequence.

ACCESSION AY109355

VERSION AY109355.1 GI:21213015

KEYWORDS HTC.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 3374)

Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,

Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.

Maize Mapping Project/DuPont Consensus Sequences for Design of

Overgo Probes

Unpublished (2002)

2 (bases 1 to 3374)

Coe, E.H.

Direct Submission

Submitted (25-APR-2002) Maize Mapping Project, University of

Missouri, Columbia, MO 65211, USA

If you are interested in getting corresponding physical clones,

these are publicly available from ZmDB and may be found by BLAST

searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,

www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the

maize cDNA sequences is either Virginia Walbot, Stanford or Pat

Schnable, Iowa State, then clones may be requested from ZmDB:

www.zmdb.iastate.edu.

Location/Qualifiers

1. .3374

/organism="Zea mays"

FEATURES

source

ORIGIN

Query Match 41.6%; Score 559; DB 13; Length 722;

Best Local Similarity 90.9%; Pred. No. 4.6e-76;

Matches 639; Conservative 0; Mismatches 60; Indels 4; Gaps 4;

QY 37 CCCCTTCCCATACCATGTCGAGCTTCCACCGCGGAGCACCAGTCCGCGGCACCGC 96

Db 16 CTTCTCGCCCTCACCATGTCGAGCTTCCACCGCGGAGCACCAGTCCGCGGCACCGC 74

QY 97 GCTCGGCGAGAGTGGCGGCGCTCATCGAGCGCTCGGCGCTTCTTACAGCGGCTC 156

Db 75 GCTTCGGCGCAACAGCTGGGTTCGCTCATCGAGCGCTTCTTACAGCGGCTC 134

QY 157 CAGCGCGCGGACCGTGGGAGCAGCAGGTCGCGCTTCTATAGGCGTTCCTCCGCCACCGC 216

Db	62	ATCCGGGACACCTTCTTCCCGGGGTTTACGGACAGGACTCTCTCCCAACCGAGGGCGGG	121
QY	292	CCCGGGAGCCGATCCTCACTCGTCTCGACGACCTCTCTCGGGGTTTTCAGGCGCC	351
Db	122	CCAGGGGAGCCGATCCGCACCTCGTCTCGACGACCTCTCTCGGGGCTTGGAGGCGCC	181
QY	352	TGCGTCGACAGATCAAGATCGGCGCCATACGTFGGCCACGAGTTTCGCCGAGCCCTAC	411
Db	182	TGCGTCGCGGACATCAAGATCGGCGCCATACGTFGGCCGCGAGCTCGCGGAGCCCTAC	241
QY	412	ATTCGCAAGTGCTCCGCAAGGACCGGGGACCAACGAGCGTTCTGTTCGATTTCCGCGTC	471
Db	242	GTTCGCCAGTGCTCCGCAAGACCGCGGACACAGAGCAITCTGCTCGAATTCGCGTC	301
QY	472	TCCGGGTCGAGTCTGTCCGCCCGCGAGGGCGCGTGTGGCGGACGAGCGCCCGAGGTG	531
Db	302	TCCGGGTCGCGGTCGTCCGTCCCGAGGGCGCGTGTGGCGGACTGAGCGCCCGAGGTG	361
QY	532	AAGGCCATGGACACCGCGCGGTCCGCGCGGTCTCCGGCGCTACGTGTCACTCCGTGCC	591
Db	362	AAGGCCCTGGACACCGCGCGGTCCGCGCGGTCTCCGGCGCTACGTTCATCCGTGCC	421
QY	592	GACGAGGGATGGACTGTGCGCTCCCGCGCGCGGTGTACGGAGGAAAAGTGGAGTCTTG	651
Db	422	GACGAGGGATGGACTGTGCGCACTCCCGCGCGGTGTACCGCGCGAGGGGAGTCTTG	481
QY	652	TCAAGCTGCGGAGCTCAAGGCTGGTTTCAGGAGCAGACTGTGTCCACTTCTACTCG	711
Db	482	TCAAGCTGCGGAGCTTAAAGCGTGGTTTCAGGAGCAGACTCTGTCTCCACTTCTACTCG	541
QY	712	GGTCGATCTTCTGGGCTATGATGCTGCTGCACTCGACAGCGGAGATGGGGTGG	771
Db	542	GCGGTGAATCTTTTGGGCTATGATGCTAGTGCAGTGCACAGCGGAGGTGAAGTGG	601
QY	772	GTGACGCTGAAGCTGGTGGACTTTTCCCATGTGCGCCGAGGGTGATGGGGTGATTGACCAC	831
Db	602	TTGAGGTGAA-CTTGTTGACTTTTCCCATTTGTCCAGTTGATAGGGGGATGAACCA	660
QY	832	AACTTCCTGGG	842
Db	661	AAATTTCTGGG	671

RESULT 13	
CC724950/c	
LOCUS	
DEFINITION	CC724950 OGLAB23TH ZM 0.7_1.5 KB Zea mays genomic clone ZMWBMa0307C22, genomic survey sequence.
ACCESSION	CC724950
VERSION	CC724950.1 GI:32143883
KEYWORDS	GSS.
SOURCE	Zea mays
ORGANISM	Zea mays
REFERENCE	Bakayota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 722)
AUTHORS	Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nurnberg,A., Robbins,D. and Lakey,N. Consortium for Maize Genomics Unpublished (2002) Other GSSs: OGLAB23TV Contact: Cathy Whitelaw
TITLE	TIGR
JOURNAL	9712 Medical Center Drive, Rockville, MD 20850, USA
COMMENT	Tel: 301-838-5843 Fax: 301-838-0208 Email: whitelaw@tigr.org Seq primer: TR Class: sheared ends. Location/Qualifiers 1..722 source
FEATURES	

```
/organism="Zea mays"  
/mol_type="genomic DNA"  
/strain="B73"  
/db_xref="taxon:4577"  
/clone="ZMMEMa0307C22"  
/clone_lib="ZM_0.7_1.5_KB"  
/note="Vector: pBGSK-; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"
```

ORIGIN

Query Match	40.2%;	Score 540.8;	DB 29;	Length 722;
Best local Similarity	89.6%;	Pred. No. 2.8e-73;		
Matches 638;	Conservative 0;	Mismatches 17;	Indels 57;	Gaps 3;
<hr/>				
QY	246	CTTCCCCCGGTTCCACGGCACGGACTCTCTCCCCACCGAGGCGCAGCCGGGAGCGCA	305	
DB	722	CTTCCCCCGGTTCCACGGCACGGACTCTCTCCACCGAGGCGCAGCCGGGAGCGCA	663	
QY	306	TCTTCACTCTGCTCTCGAGACCTCTCTCGGGGGTTTCAAGCGCCCTGTGGTCGACACAT	365	
DB	662	TCCGCACCTCTGCTCTCGAGACCTCTCTCGGGGGTTTGAAGCGCCCTGTGGTCGACACAT	603	
QY	366	CAAGATCGGCGGCATCACGTTGGCCACCGAGTTTCGCCGGAGCCCTACATCGCCAAAGTGCCT	425	
DB	602	CAAGATCGGTCGCACTCAAGT-----	583	
QY	426	CGCCAGGACCGGGGACACAGAGGTTCCTGCTCGAGATTCGGGTCTCGGCGTCCGAGT	485	
DB	582	-----GACCACGAGCGTTCTGTCTCGGATTCGGGTCTCGGCGTCCGAGT	538	
QY	486	CGTCGGCCCCGAGGGCGCGTGTGGCGACGAGCGCCCGGAGGTCAAGGCCCATGACAC	545	
DB	537	CGTCGCCCCGAGGGCGCGTGTGGCGACGAGCGCCCGGAGGTCAAGGCTATGACAT	478	
QY	546	CGCCGGCGTCCGGCGGTGCTTCGGCGCTACGTGTATCCG-TTGGCGACGAGGGGATGG	604	
DB	477	TGTTCGGCGTCCGGCGGTGCTTCGGCGCTACGTGTATCCGCTTTCGCGACGAGGGGATGG	418	
QY	605	ACTGTTCGCTTCGCCCGGGCGGTGTACGGAGGAAAAGGTGGAGTCTTGTGTACACGCTCGCG	664	
DB	417	ACTGCGCGTTCGCCCGGGCGGTGTACGGAGGAAAAGGTGGAGTCTTGTGTACACGCTCGCG	358	
QY	665	AGCTCAAGCGGTGTTTCGAGGAGCAGACTGTGTTCCATTCTACTCGGCGTGCATTCTTC	724	

RESULT 14
RZ774817

BZ774817	BZ774817	593 bp	DNA	linear	GSS 14-MAR-2003
LOCUS	i151a11.b1	WGS-Zmaysf	(DH5a methyl filtered)	zea mays	genomic clone
DEFINITION	i151a11	genomic survey sequence.			
ACCESSION	BZ774817				
VERSION	BZ774817.1	GI:28952260			
KEYWORDS	GSS.				
SOURCE	zea mays				
ORGANISM	zea mays				

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 593)

Rabinowicz, P.D., O'Shaughnessy, A.L., Balija, V., Dedhia, N., Katzenburger, F., King, L., Miller, B., Muller, S., Nascimento, L., Zutavern, T., McCombie, W.R. and Martienssen, R.A.

Genomic shotgun sequences from Zea mays (methyl-filtered)

Unpublished (2002)

Contact: W. Richard McCombie

Lita Annenberg Hazen Genome Sequencing Center

Cold Spring Harbor Laboratory

PO Box 100, Cold Spring Harbor, NY 11724, USA

Tel: 516 367 8884

Fax: 516 367 8874

Email: mcombie@cshl.org

Plate: i151 row: a column: 11

Seq primer: -21M3UnivFwd

Class: shotgun

High quality sequence stop: 593.

Location/Qualifiers

1. 593

/organism="Zea mays"

/mol_type="genomic DNA"

/cultivar="B73"

/db_xref="taxon:4577"

/clone="i151a11"

/lab_host="DH5a"

/clone_lib="WGS-ZmaysF (DH5a methyl filtered)"

/note="Organ: immature ears; Site 1: Xba I; Site 2: Xba I; The vector was digested with XbaI and one nucleotide was added by fill in in the recessive 3' end. The genomic DNA was nebulized, end repaired, adaptor ligated and size fractionated using sephadex. The resulting fragments were between 0.8 and 3 kb and were cloned into the vector (.x/y reads in M13mp19, .b/g reads in pUC19). The same ligation was transformed into DH5a."

FEATURES

source

Location/Qualifiers

1. 593

/organism="Zea mays"

/mol_type="genomic DNA"

/cultivar="B73"

/db_xref="taxon:4577"

/clone="i151a11"

/lab_host="DH5a"

/clone_lib="WGS-ZmaysF (DH5a methyl filtered)"

/note="Organ: immature ears; Site 1: Xba I; Site 2: Xba I; The vector was digested with XbaI and one nucleotide was added by fill in in the recessive 3' end. The genomic DNA was nebulized, end repaired, adaptor ligated and size fractionated using sephadex. The resulting fragments were between 0.8 and 3 kb and were cloned into the vector (.x/y reads in M13mp19, .b/g reads in pUC19). The same ligation was transformed into DH5a."

ORIGIN

Query Match 39.9%; Score 535.6; DB 28; Length 593;
Best Local Similarity 97.9%; Pred. No. 1.8e-72;
Matches 564; Conservative 0; Mismatches 9; Indels 3; Gaps 2;

QY 42 TCCCATACCATGTCGACCTCCACCGCCGAGCACCACCAAGTCGCGGCCACCGGCCTC 101

DB 20 TCCCATACCATGTCGACCTCCACCGCCGAGCACCACCAAGTCGCGGCCACCGGCCTC 79

QY 102 CGCAGCAAGCTGGCCCGCTCATCGAGCTCCGGCTCTTCTACAGCGCTCCAGGC 161

DB 80 CGCAGCAAGCTGGCCCGCTCATCGAGCTCCGGCTCTTCTACAGCGCTCCAGGC 139

QY 162 CGCAGCAAGCTGGCCCGCTCATCGAGCTCCGGCTCTTCTACAGCGCTCCAGGC 221

DB 140 CGCAGCAAGCTGGCCCGCTCATCGAGCTCCGGCTCTTCTACAGCGCTCCAGGC 199

QY 222 CGCAGCAAGCTGGCCCGCTCATCGAGCTCCGGCTCTTCTACAGCGCTCCAGGC 281

DB 200 CGCAGCAAGCTGGCCCGCTCATCGAGCTCCGGCTCTTCTACAGCGCTCCAGGC 259

QY 282 CGCAGCAAGCTGGCCCGCTCATCGAGCTCCGGCTCTTCTACAGCGCTCCAGGC 341

DB 260 CGCAGCAAGCTGGCCCGCTCATCGAGCTCCGGCTCTTCTACAGCGCTCCAGGC 319

QY 342 TCAGCGCCCTGCTGTCGAGACATCAAGATGGGCCATACGTCGGCCACCGAGTTCGCC 401

DB 320 TGAGCGCCCTGCTGTCGAGACATCAAGATGGGCCATACGTCGGCCACCGAGTTCGCC 379

QY 402 GGAGCCCTACATCGCCCAAGTGCCTCGCCAAAGACCGCGGGACCAAGAGCTTCTGCTCGG 461

DB 380 GGAGCCCTACATCGCCCAAGTGCCTCGCCAAAGACCGCGGGACCAAGAGCTTCTGCTCGG 439

QY 462 ATTCGCGCTCTCGCGCTTCGAGTCTCGGCCCGCGAGCGCGCGTGTGGCGGACGAGCGC 521

Db 440 ATTCGCGCTCT--TCGTCCTCGAGTCGTCGGCCCCGAGGGCGCGTGTGCGGACGAGCG 497

QY 522 CCCGAGGTGAAGCCATGAGACACCGCGCGGTTCGCGCGGTGCTTCGCGCGCTACGTGTC 581

DB 498 CCCGAGGTGAAGCCATGAGACACCGTCGGGTTCGCGCGGTGCTTCGCGCGCTACGTGTC 557

QY 582 ATTCGCTTCGCGACGAGGGGATGACTGTGCGCTCG 616

DB 558 ATTCGCTTCGCGACGAGGGGATGACTGTGCGCTCG 593

RESULT 15

CA130686

LOCUS

DEFINITION

5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Saccharum officinarum

Saccharum officinarum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 744)

The libraries that made SUCEST

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

Contact: Arruda P

Centro de Biologia Molecular e Engenharia Genetica

Universidade Estadual de Campinas

Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137

Fax: 55 19 3788 1089

Email: parruda@unicamp.br

Clone distribution: clone distribution information can be found

through the Brazilian Clone Collection Center (BCCC) at

http://www.bcccenter.fcav.unesp.br

Plate: 004 row: G column: 06

Seq primer: T7 Promoter Primer.

Location/Qualifiers

1. 744

/organism="Saccharum officinarum"

/mol_type="mRNA"

/db_xref="taxon:4547"

/clone="SCCRL1004G06"

/lab_host="DH10B"

/clone_lib="RT1"

/note="Organ: Root tips (0.3cm-long) from adult plants; Vector: pSport1; Site 1: SalI; Site 2: NotI; An

unidirectional cDNA library generated from [Root tips

(0.3cm-long) from adult plants]. cDNA was prepared from

polyA+ mRNA using SuperScript Plasmid System kit

(Invitrogen). The double-strand cDNAs were fractionated

in a sepharose CL-2B 40cm-columns and fragments sizing

between 0.8 and 1.5 Kb were directionally cloned into the

vector. Details of each source of RNA and library

construction can be obtained at

http://sucest.lad.ic.unicamp.br/public"

ORIGIN

Query Match 39.1%; Score 525.6; DB 13; Length 744;

Best Local Similarity 88.6%; Pred. No. 6e-71;

Matches 592; Conservative 0; Mismatches 74; Indels 2; Gaps 2;

QY 37 CCCCTTCCCATACCATGTCGACCTCCACCGCCGAGCACCACCAAGTCGCGGCCACCGC 96

DB 16 CCTCTGCGCTTCCCATGTCGACCT-CGCCCCCGGAGCACCACCAAGTCGCGGCCACCGC 74

QY 97 GCCTCCGCCAGCAAGCTGGGCCCGCTCATCGACGGCTCCGGCTCTTCTTACAGCGGCTC 156

DB 75 GCCTCCGCCAACAGCTGGGTTCGCTCATCGACGGCTCTGGCCCTTCTTACAGCGGCTC 134

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2004, 23:22:54 ; Search time 5348 Seconds
(without alignments)
10892.490 Million cell updates/sec

Title: US-10-042-894A-7

Perfect score: 1344

Sequence: 1 gcacgagtcagtcgctcac.....ataaaaaaaaaaaaaaaa 1344

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_man.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1344	100.0	1344	6	AX513570	AX513570 Sequence
2	865.4	64.4	923	6	AX513568	AX513568 Sequence
3	844.8	62.9	1169	6	AX513564	AX513564 Sequence
4	837.4	62.3	923	6	AX513566	AX513566 Sequence
5	740.4	55.1	3416	6	AX513583	AX513583 Sequence
6	686	51.0	899	6	AX513578	AX513578 Sequence
c 7	546.4	40.7	132170	8	AP005749	AP005749 Oryza sat
8	544.8	40.5	1570	8	AK072296	AK072296 Oryza sat
9	470.4	35.0	643	6	AX513580	AX513580 Sequence
10	298.4	22.2	519	6	AX513581	AX513581 Sequence
11	191.6	14.3	353	6	AX513582	AX513582 Sequence
12	184.6	13.7	1195	6	AX513574	AX513574 Sequence
13	181.6	13.5	25054	8	LUS310150	AXJ10150 Linum usi
14	179	13.3	1105	6	AX513572	AX513572 Sequence
15	156.2	11.6	116893	8	AP004772	AP004772 Oryza sat
16	149.8	11.1	1020	6	AX513576	AX513576 Sequence
17	148.6	11.1	1390	8	AY136378	AY136378 Arabidops
18	148.6	11.1	116783	8	ATT211	AL163912 Arabidops
19	148.2	11.0	905	8	BT000196	BT000196 Arabidops
20	147.8	11.0	115040	8	AC124954	AC124954 Medicago
21	146.6	10.9	1243	8	AY087217	AY087217 Arabidops
22	145.8	10.8	129248	2	AC127018	AC127018 Medicago
c 23	145.6	10.8	903	8	ATH243592	AJ243592 Arabidops
24	145.6	10.8	903	8	AY072621	AY072621 Arabidops
25	145.6	10.8	903	8	AY147936	AY147936 Arabidops
26	145.6	10.8	1166	8	AF412073	AF412073 Arabidops
27	145.6	10.8	1170	8	ATH245521	AJ245521 Arabidops
c 28	145.6	10.8	57246	8	AB010069	AB010069 Arabidops
29	145.4	10.8	861	8	ATH404678	AJ404678 Arabidops
30	145.4	10.8	861	8	AY147935	AY147935 Arabidops
31	143.6	10.7	1130	8	AY085862	AY085862 Arabidops
c 32	83.8	6.2	43428	8	AP004981	AP004981 Lotus cor
33	75.2	5.6	110000	2	LMFLCHR32_06	Continuation (7 of
34	75	5.6	110000	2	LMFLCHR36_03	Continuation (4 of
35	74.8	5.6	296500	1	SCO9319128	AL9319128 Streptomy
36	73.8	5.5	125020	9	AF429315	AF429315 Homo sapi
c 37	73.6	5.5	110000	2	LMFLCHR36_31	Continuation (32 o
38	73.4	5.5	300800	1	SCO939112	AL939112 Streptomy
39	72.4	5.4	113193	1	AF357202	AF357202 Streptomy
40	72.4	5.4	113193	6	AX703543	AX703543 Sequence
c 41	71.4	5.3	135638	1	AF484556	AF484556 Streptomy
42	70.4	5.2	224	8	AY022706	AY022706 Oryza sat
c 43	70.4	5.2	78220	2	AC023212	AC023212 Homo sapi
44	69.6	5.2	1926	6	AR217867	AR217867 Sequence
45	69.6	5.2	1931	6	AR083152	AR083152 Sequence

ALIGNMENTS

RESULT 1	AX513570	AX513570	1344 bp	DNA	linear	PAT 05-OCT-2002
LOCUS	AX513570	Sequence 7 from Patent WO02059324.				
DEFINITION	AX513570					
ACCESSION	AX513570					
VERSION	AX513570.1	GI:23559670				
KEYWORDS						
SOURCE	Zea mays					
ORGANISM	Zea mays					
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD					
	clade; Panicoideae; Andropogoneae; Zea.					
REFERENCE	1					
AUTHORS	Shi,J., Beach,L.R., Wang,H., Rafalski,J.A. and Cahoon,R.E.					
TITLE	Novel inositol polyphosphate kinase genes and uses thereof					

Pred. No. is the number of results predicted by chance to have a

/translation="MPDLHPPEHOVAGHRASKEPGLIDGSLFYKPLQAGDRGEHE
VAFYEAFSAHAAPARIDTTFEPHGTLLPTEAPGCEPHPHVLVDLLAGFAPCV
ADIKIGALTWPSPPEPIAKCLMDRGTTSVLLGFRVSGVRVPEGAVWRTERPEV
KAMDTAGVRRVLRVSVSVADEGMDCALAAAVYGGKGVLSQLRELKAWFEQTLFHF
YSASILIGYDAAAVAAAGGGGVTVKLVDFAHVAAGDGVIDHNFILGLCSLIFKVS
VPEPQTQPLGFS"

ORIGIN

```
Query Match      64.4%; Score 865.4; DB 6; Length 923;
Best Local Similarity 98.8%; Pred. No. 5.5e-121;
Matches 870; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 42 TCCCATACCATGTCGACCTCCACCCGCGAGCACCAAGTCGCGCCACCGCGCTC 101
DB 43 TCCCATACCATGTCGACCTCCACCCGCGAGCACCAAGTCGCGCCACCGCGCTC 102
QY 102 CGCAGCAAGTGGGCCCGCTCATCGAGCGTTCGCGCTCTTCTACAAAGCGCTCCAGGC 161
DB 103 CGCCAGCAAGCGCGCGCTCATCGAGCGTTCGCGCTCTTCTACAAAGCGCTCCAGGC 162
QY 162 CGCGACCGTGGGAGCAGAGTGCCTTCTATGAGCGTTCGCGCCACCGCGCTC 221
DB 163 CGCGACCGTGGGAGCAGAGTGCCTTCTATGAGCGTTCGCGCCACCGCGCTC 222
QY 222 CGCGCCCGCATCCGAGACACTTCTCCCGGTTCCAGCGCAGCGACTCCTCCCCAC 281
DB 223 CGCGCCCGCATCCGAGACACTTCTCCCGGTTCCAGCGCAGCGACTCCTCCCCAC 282
QY 282 CGAGCGCAGCCGCGGAGCGCATCTCTACCTCGTCTCGACGACTCTCTCGGGGTT 341
DB 283 CGAGCGCAGCCGCGGAGCGCATCTCGACCTCGTCTCGACGACTCTCTCGGGGTT 342
QY 342 TCAGCGCCCTCGCTCGCAGACATCAAGATCGGCGCATACAGTGGCCACCGAGTTCGCC 401
DB 343 TGAGCGCCCTCGCTCGCAGACATCAAGATCGGCGCATACAGTGGCCACCGAGTTCGCC 402
QY 402 GGAGCCCTACATCGCCAAAGTGCCTCGCCAGAGCCGCGGACCAAGCGTTCCTCGG 461
DB 403 GGAGCCCTACATCGCCAAAGTGCCTCGCCATGACCGCGGACCAAGCGTTCCTCGG 462
QY 462 ATTCGCGCTCTCGCGCTCCGAGTCTCGGCGCCGAGCGCGCTGTGGCGGACGAGCG 521
DB 463 ATTCGCGCTCTCGCGCTCCGAGTCTCGGCGCCGAGCGCGCTGTGGCGGACGAGCG 522
QY 522 CCCGAGGTGAAGCCATGACACCGCGCGGTTCGCGCGGTCTCGCGCGCTACGTTGC 581
DB 523 CCCGAGGTGAAGCCATGACACCGCGCGGTTCGCGCGGTCTCGCGCGCTACGTTGC 582
QY 582 ATCCGTTGCCAGCAGGAGGATGAGTGTGCGCTCGCGCGCGGTGTACGAGGAAAGG 641
DB 583 ATCCGTTGCCAGCAGGAGGATGAGTGTGCGCTCGCGCGCGGTGTACGAGGAAAGG 642
QY 642 TGGAGTCTTGTACAGCTGCGCAGCTCAAGCGCTGTTCGAGGAGCAGACTCTGTTCCA 701
DB 643 TGGAGTCTTGTACAGCTGCGCAGCTCAAGCGCTGTTCGAGGAGCAGACTCTGTTCCA 702
QY 702 CTTTACTCTCGCGTGCATCTTCTTGGGCTATGATGCTGCTCAGTCCGAGCAGCGGAGA 761
DB 703 CTTTACTCTCGCGTGCATCTTCTTGGGCTATGATGCTGCTCAGTCCGAGCAGCGGAGG 762
QY 762 TGGGGTGGGGTGAGCGTGAAGCTGGTGAATTTTGCCCATCTGCGCGAGGCTGATGGGT 821
DB 763 TGGGGTGGGGTGAGCGTGAAGCTGGTGAATTTTGCCCATCTGCGCGAGGCTGATGGGT 822
QY 822 GATTGACACAACTCTCTCGCGCGGCTCTGCTCGCTGATCAAGTTCGTTTCTGACATGT 881
DB 823 GATTGACACAACTCTCTCGCGCGGCTCTGCTCGCTGATCAAGTTCGTTTCTGACATGT 882
QY 882 TCCGAGACTCTCTATACGAGCCTTTGGGTCCTTCTTAAG 922
DB 883 TCCGAGACTCTCTCAGACGACCTTTGGGTCCTTCTTAAG 923
```

RESULT 3

AX513564 LOCUS 1169 bp DNA linear PAT 05-OCT-2002
DEFINITION Sequence 1 from Patent WO2059324.
ACCESSION AX513564
VERSION AX513564.1 GI:23559664
KEYWORDS

SOURCE

Zea mays
ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

1 Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.

Novel inositol polyphosphate kinase genes and uses thereof

Patent: WO 02059324-A 1 01-AUG-2002;

PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES

Location/Qualifiers
source

1..1169

/organism="Zea mays"

/mol_type="unassigned DNA"

/db_xref="taxon:4577"

84..806

/note="unnamed protein product"

/codon_start=1

/protein_id="CAD52951.1"

/db_xref="GI:23559665"

/db_xref="RENTREMBL:CAD52951"

/translation="MPDLHPPEHOVAGHRASKEPGLIDGSLFYKPLQAGDRGEHE

VAFYEAFSAHAAPARIDTTFEPHGTLLPTEAPGCEPHPHVLVDLLAGFAPCV

ADIKIGALTWPSPPEPIAKCLMDRGTTSVLLGFRVSGVRVPEGAVWRTERPEV

KAMDTAGVRRVLRVSVSVADEGMDCALAAAVYGGKGVLSQLRELKAWFEQTLFHF

YSASILIGYDAAAVAAAGGGGVTVKLVDFAHVAAGDGVIDHNFILGLCSLIFKVS

VPEPQTQPLGFS"

ORIGIN

```
Query Match      62.9%; Score 844.8; DB 6; Length 1169;
Best Local Similarity 97.2%; Pred. No. 6.8e-118;
Matches 890; Conservative 0; Mismatches 22; Indels 4; Gaps 3;

QY 42 TCCCATACCATGTCGACCTCCACCCGCGGAGCACCAAGTCGCGCCACCGCGCTC 101
DB 74 TCCCATACCATGTCGCGCTCCACCCGCGGAGCACCAAGTCGCGCTC 133
QY 102 CGCAGCAAGTGGGCCCGCTCATCGAGCGTTCGCGCTCTTCTACAAAGCGCTCCAGGC 161
DB 134 CGCAGCAAGTGGGCCCGCTCATCGAGCGTTCGCGCTCTTCTACAAAGCGCTCCAGGC 193
QY 162 CGCGACCGTGGGAGCAGAGTTCGCTTATGAGCGTTCGCGCCACCGCGCT 221
DB 194 CGCGACCGTGGGAGCAGAGTTCGCTTATGAGCGTTCGCGCCACCGCGCT 253
QY 222 CGCGCCCGCATCCGAGACACTTCTTCCCGCGGTTCACGCGACGCGACTCTCTCCAC 281
DB 254 CGCGCCCGCATCCGAGACACTTCTTCCCGCGGTTCACGCGACGCGACTCTCTCCAC 313
QY 282 CGAGCGCGACCGCGGAGCGCATCTCCTGCTCCTCGAAGCCTCTCTCGCGGGTT 341
DB 314 CGAGCGCGACCGCGGAGCGCATCTCGCACCTCTCTCGACACCTCTCTCGCGGGTT 373
QY 342 TCAGGCGCTGCTGCTCGCAGACATCAAGATCGCGCCATCAAGTGGCCACCGAGTTCGCC 401
DB 374 TGAGGCGCTGCTGCTCGCAGACATCAAGATCGCGCCATCAAGTGGCCACCGAGTTCGCC 433
QY 402 GGAGCGCTACATCGCCAAAGTGCCTCGCCAAAGGACCGCGGACCAAGAGCGTTCGTCGG 461
DB 434 GGAGCGCTACATCGCCAAAGTGCCTCGCCAAAGGACCGCGGACCAAGAGCGTTCGTCGG 493
QY 462 ATTCGCGCTCTCGCGCTCGAGTTCGCGCCCGCGAGGGCGCTGTGCGGAGCAGGAGCG 521
DB 494 ATTCGCGCTCT--TGCCTCGAGTCTGCTCGGCCCGCGAGGGCGCTGTGCGGAGCAGGAGCG 551
QY 522 CCCGAGGTTGAAGGCGCATGACACCGCGCGTTCGCGCGCTGCTCGCGCGCTACGTCGTC 581
```


QY	781	AAGCTGTGGACTTTGCCCATGTGGCCGAGGGTGATGGGTGATTGACACAACTTCCTG	840
Db	747	AAGCTGTGGACTTTGCCCATGTGGCCGAGGGTGATGGGTGATTGACACAACTTCCTG	806
QY	841	GGCGGGCTCTGCTCGCTGATCAAGTTCTGTTCTGACATTGTTCCGAGACTCCTCATACG	900
Db	807	GGCGGGCTCTGCTAGCTGATCAAGTTGTTTCTGACATTGTTCCAGAGACTCCTCAGACG	866
QY	901	CAGCCTTTGGTCTCTTTTAAGACAGACTCTGSCA-TTTTCGATTTGATAACAAG	955
Db	867	CAGCCTTTGGTCTCTTTTAAGACAGACTCTGSCATTTTCGATTTGATAACAAG	922
RESULT 6			
AX513578			
LOCUS	AX513578	899 bp	DNA linear PAT 05-OCT-2000
DEFINITION	Sequence 15 from Patent WO02059324.		
ACCESSION	AX513578		
VERSION	AX513578.1	GI:23559679	
KEYWORDS	Zea mays		
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE	1 Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.		
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
TITLE	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD		
JOURNAL	clade; Panicoideae; Andropogoneae; Zea.		
FEATURES	1 .899		
source	/organism="Zea mays"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:4577"		
	89. .424		
	/notes="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAD52958.1"		
	/db_xref="GI:23559680"		
	/db_xref="RENTREMBL:CAD52958"		
	/translations="MPDLHPHQVAGHRASAKLGLIDSGLFYFPLOAGDRGEHCV		
	VAFYEAFSAHAAPVPIRDTFFPRFHGTRLLPTEAQPGEHPHVLVDLLAGFEAPCV		
	ADIKIGAIT"		
ORIGIN			
Query Match	51.0%;	Score 686;	DB 6; Length 899;
Best Local Similarity	88.7%;	Pred. No. 5.9e-94;	
Matches	802;	Conservative 0;	Mismatches 40; Indels 62; Gaps 3;
QY	2	CACGAGTGCTAGTCGGTCACCCCTCGGCCCATAGTCCCTTCCCATACCATGTCGGACC	61
Db	45	CACCATCGCCACCCGTCACCCCTTGCTCCATAG-----TCGCCATACCATGCCCCGACC	98
QY	62	TCCACCGCCGCGAGCACCAAGTCGCGGCCACCGCGCTCCGCCAGCAAGCTGGGCGCCGC	121
Db	99	TCCACCGCCGCGAGCACCAAGTCGCGGTCACCGGCTCCGCCAGCAAGCTGGGCGCCAC	158
QY	122	TCAATCAGCGCTCCGGCTCTTCTACAAGCCGCTCCAGCGCGCGACCGTGGGAGGACACG	181
Db	159	TCAATCAGCGCTCTGGCTCTTCTACAAGCCGCTCCAGCGCGCGACCGTGGGAGGACACG	218
QY	182	AGGTGCGCTTCTATAGGCGTTCTCGGCCACCGCGCGCTCCGGCCCGCATCCGAGACA	241
Db	219	AGGTGCGCTTCTATAGGCGTTCTCGGCCACCGCGCGCTCCGGCCCGCATCCGAGACA	278
QY	242	CGTTCTTCCCGGGTTCACCGCACCGGACTCTCTCCCCACCGAGGCGCAGCCCGGGGAGC	301
Db	279	CGTTCTTCCCGGGTTCACCGCACCGGACTCTCTCCCAACGAGGCGCAGCCCGGGGAGC	338
QY	302	CGCATCCTCACCTCGTCTCTCGAGCACTCTCTCGGGGGTTTTCAGCGCCCTCGCTCGCAG	361

```
Db 339 CGCATCCGTACCTGCTCTCGACGACCTCTCTCGGGGGTTTGAGGGCGCCCTGCGTCGAG 398
QY 362 ACATCAAGATCGGCGCCATCATCGTGCCACCACAGTTTCGCGGAGCCCTACATCGCCAAAGT 421
Db 399 ACATCAAGATCGGTCCCATCAGT----- 422
QY 422 GCCTGCCCAAGACCGCGGGACCGAGCGTTCTGCTCGGATTCGGGTCTCCGCGCTCC 481
Db 423 -----GACCATGAGCGATCTGCTCGGATTCACGCTCTCCGCGCTCC 463
QY 482 GAGTCTGTCGCGCCCGAGGGCGCGTGTGGCGGACCGAGCGCCGCGAGGTGAAGGCCATGG 541
Db 464 GAGTCTGTCGCGCCCGAGGGCGCGTGTGGCGGACCGAGCGCCCTGAGGTGAAGGCTATGG 523
QY 542 ACACCGCGGGTCTCGCGCGCGTCTCCGCGGCTACGTGTCAFCG- TTGCCGACGAGGGG 600
Db 524 ACATTGTGCGGCTCCGCGCGCTGCTCCGCGGTGTCATGTCATCCGCTTGCGCGCGAGGGG 583
QY 601 ATGGACTGTGCGTCCGCGCGCGGTGTACGAGGAGAGTGGAGTCTTGTCAAGCTG 660
Db 584 ATGGACTGTGCGTCCGCGCGCGGTGTACGAGGAGAGTGGAGTCTTGTCAAGCTG 643
QY 661 CCGAGACTCAAGCGCTGTTGAGGAGCAGACTCTGTTCCACTTCTACTCGCGCTCGATT 720
Db 644 CCGAGACTCAAGCGTGGTTCGAGGGGAGACTCTGTTCCACTTCTACTCGCGCTCGATT 703
QY 721 CTTCTGGGCTATGATGCTGCTGACGTGCGAGCGAGTGTGGGGTGGGGTGAAGCTG 780
Db 704 CTTCTGGGCTATGATGCTGCTGACGTGCGAGCGAGGTGGGGTGAAGCTG 763
QY 781 AAGCTGGTGGACTTGGCCATGTCGCGGAGGTGATGGGTGATGACCAACACTTCCTG 840
Db 764 AAGCTGGTGGACTTGGCCATGTCGCGGAGGTGATGGGTGATGACCAACACTTCCTG 823
QY 841 GCGGGGCTGCTGCTGATCAAGTTCGTTCTGACATTTGTCGAGAGACTCTCATACG 900
Db 824 GCGGGGCTGCTGATGATCAAGTTTGTCTGACATTTGTCGAGAGACTCTCATACG 883
QY 901 CAGC 904
Db 884 CAGC 887

RESULT 7
AP005749/c
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,
BAC clone:OSUNBa0047A17, complete sequence.
ACCESSION
AP005749
VERSION
AP005749.2 GI:34740257
KEYWORDS
HTG.
SOURCE
Oryza sativa (japonica cultivar-group)
ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
1
Sasaki,T., Matsumoto,T. and Katayose,Y.
Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC
clone:OSUNBa0047A17
Published Only in Database (2002)
JOURNAL
2 (bases 1 to 132170)
Sasaki,T., Matsumoto,T. and Katayose,Y.
Direct Submission
Submitted (18-SEP-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program, Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasakia@nias.affrc.go.jp, URL:http://rsp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
On Sep 15, 2003 this sequence version replaced gi:23200608.
The orientation of the sequence is from M13rev to -21M13 of the BAC
clone.
FEATURES
Location/Qualifiers
```

```
source
1. 132170
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="2"
/clone="OSUNBa0047A17"

ORIGIN
Query Match 40.7%; Score 546.4; DB 8; Length 132170;
Best Local Similarity 77.5%; Pred. No. 3.1e-73;
Matches 718; Conservative 0; Mismatches 191; Indels 17; Gaps 4;

QY 13 TCCGTCACCCCTCGCGGCCCATAGTCCCTTCCCATACCATGTCGCGACCTCCACCCGCG 72
Db 1983 TCCACCAACCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1926
QY 73 GAGCACCAAGTCGCGGCCACCGCGCGCTCCGCGCAGCAAGCTGGCGCCGCTCATGACGGC 132
Db 1925 GAGCACCAAGTCGCGGCCACCGCGCGCTCCGCGCAGCAAGCTGGCGCCGCTCATGACGGC 1866
QY 133 TCCGGCTCTTCTAACAAGCCGCTCCAGCCCGGACCGCTGGGGAGCAGAGGTGCGCTTC 192
Db 1865 GAGGGGCTCTTCTAACAAGCCGCTCCAGCCCGGAGCGCGGGAGCAGAGGCGCGCTTC 1806
QY 193 TATGAGGGTCTCGCGGCCACCGCGCGCTCCGCGCGCGCATCCCGAGACACCTTCTTCCC 252
Db 1805 TACCGCGGTTTACCGCGCGCACCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCG 1746
QY 253 CGGTTCCACGGCAGCGACTCTCTCCCAACGAGCGCAGCGC---CGGGAGCCGATCCT 309
Db 1745 CGTTCACGGCAGCGCTTCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1686
QY 310 CACTGTGTCGCGAGCACTCTCTCGGGGGTTTCAGGGGCCCTCGCTCGCAGACATCAAG 369
Db 1685 CACATCGTCTCGAGCACTCTCTCGGGGCTCTCCGTCGCCCTCGCTCGCGCGAGCTCAAG 1626
QY 370 ATCGCGCCATCAGTGGCCACAGTTCGCGGAGCCCTACATCGCCCAAGTGCCTCGCC 429
Db 1625 ATCGCGCGCTGACGTGGCGCGCGCGATTCGCCGAGCCCTACGTCGCCCAAGTGCCTCGCC 1566
QY 430 AAGACCGCGGGACCAAGAGGTTCTGTCGGAATTCGCGCTCTCGCGCGTCCGAGTGCCTC 489
Db 1565 AAGACCGCGGAGACCAACAGCGCGCTCTCTCGGCTTCGCGCTCTCGGGGTGCTC 1506
QY 490 ---GGCCCCGAGGGCGCGCTGTGGGACAGGAGCGCCCGAGGTGAAGCCATGGACACC 546
Db 1505 GATGCCCGGGCGCGCGCTGTGGCGCGCGCGCGCTCGGAGCTGAAGGGGATCGACGCC 1446
QY 547 GCGGCGTCCGCGCGCTCTCGCGCGCTACGTGTATCCGTTGCGCAGAGGGGATGGAC 606
Db 1445 GCGGGGTTCCGCGCGCTCTCGCGCGTACGTGTATCCGCGCGCGCGCGCGCGCGCGCG 1386
QY 607 TGTGCGCTCGCGCGCGCGGTGTACGAGGAAAGGTGGAGTCTTGTCAAGCTCGCGAG 666
Db 1385 TGGCGCTCGCGCGCGCGGTGTACGAGGGGAGGGCGCGCTTGGCTCAGCTGCGGGAG 1326
QY 667 CTCAGGCGTGTTCGAGGACAGACTCTGTTCACATTTCTACTCGCGCTCGATTTCTTG 726
Db 1325 CTCAGGCGTGTTCGAGGACAAACCTGTACCACTTCTACTCGCGCTCGATTTCTTGTC 1266
QY 727 GGCT-----ATGATGCTGTGCGTCCGAGCGCGCGAGATGGGGGTGGGTGAGC 777
Db 1265 GGCTACGACGCAATGCGCGCGCGCGCTGCTCTCCGAGGTGGAGGGCGGTGAAGG 1206
QY 778 GTGAAGCTGGTGGACTTTTGGCCATGTGGCGGAGGGTGTATGGGTGATGACCAACATTC 837
Db 1205 GTGAAGCTGGTGGACTTTCGCGCATGTGACGATGGGGACGGGTGATGACCAACATTC 1146
QY 838 CTGGCGGGGCTCTGCTGCTGATCAAGTTCGTTTCTGACATTTCTCGGAGACTCTCAT 897
Db 1145 TTGGCGGGGCTCTGCTGCTCATCAAGTTCATCGCGCGACATTTGTGCGCGGTACCAG 1086
```

```

QY 898 AGCAGCCTTTGGGCTCTTAA 923
DB 1085 AAGCGCTTTCAGATCATCTTGAAA 1060

RESULT 8
LOCUS AK072296
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone J023018G11, full
insert sequence.
ACCESSION AK072296
VERSION AK072296.1 GI:32982319
KEYWORDS FLI CDNA; CAP trapper.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzae; Oryza.

1
The Rice Full-Length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-Length cDNA Project Team:
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group: Ootomo, Y., Murakami, K.,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
Kusumegi, T., Oka, M., Ryu, J., Ueda, M., Matsubara, K., RIKEN:
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
Hara, A., Hashizume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Oka, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Yoshino, M., and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
JOURNAL Science 301 (5631), 376-379 (2003)
MEDLINE 22752273
PubMed 12869764

2 (bases 1 to 1570)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayaehiraki, Y., Hayatsu, N., Hiramoto, K., Hiroaka, T.,
Hori, F., Hotta, I., Iida, Y., Iida, Y., Ikeda, R., Imamura, K.,
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
Mizuno, K., Murakami, K., Murata, J., Nagata, T., Nakamura, M.,
Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K.,
Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,
Osato, N., Oka, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,
Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and
Yoshimura, A.
Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,
Tel: 81-29-838-7007, Fax: 81-29-838-7007)
This clone is one of the 28k full-length cDNA clones from japonica
rice.
URL: http://cdna01.dna.affrc.go.jp/cdna/
NTAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K.,

FEATURES
            source
            location/Qualifiers
1..1570
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023018G11"

ORIGIN

Query Match 40.5%; Score 544.8; DB 8; Length 1570;
Best Local Similarity 77.4%; Pred. No. 1e-72;
Matches 717; Conservative 0; Mismatches 192; Indels 17; Gaps 4;

QY 13 TCGGTACCCCTTCGGCCCATAGTCCCTTCCTCCCATACCATGTCGACCTCCACCGCGC 72
DB 344 TCCACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 401

QY 73 GAGCACCAGTTCGGCGCCACCGCGCTCCGCGAGCAAGCTGGCGCCGCTCATCGAGCGC 132
DB 402 GAGCACCGAGTGGCGGGGCGACCGCGCTCCGCGACAGCTGGCGCCGCTCGTCGAGCGC 461

QY 133 TCCGGCGCTCTTCTACAGCGCGCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCTTC 192
DB 462 GAGGGGCTCTTCTACAGCGCGCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCTTC 521

QY 193 TATGAGCGCTTCGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTC 252
DB 522 TAGCGCGCGCTTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTC 581

QY 253 CGGTTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 309
DB 582 CGTTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 641

QY 310 CACCTCGTCTTCGACGACCTTCCTTCGCGGGGCTTTTCAGCGCGCGCTTCGTCGAGACATCAAG 369
DB 642 CACATCGTCTTCGACGACCTTCCTTCGCGGGGCTTCCTTCGCGCGCGCTTCGTCGAGCGTCAAG 701

QY 370 ATCGGCGCGCATACGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 429
DB 702 ATCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 761

QY 430 AAGGACCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 489
DB 762 AAGGACCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 821

QY 490 ---GGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 546
DB 822 GATGCGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 881

```

```

QY 547 GCCGCGTCCGCGCGTGTCTCCGCGCTACGTGTGTCATCGTGTGCCGACGAGGGGATGGAC 606
Db 582 GCGGGGTCCGCGCGTGTCTCCGCGCTACGTGTGCCGCGCGGCGACGGCCTGGAC 941
QY 607 TGTGCGTCCGCGCGCGTGTACGAGGAAAGTGGAGTCTTCTCAGCTCGCGAG 666
Db 942 TCGCGCTCCGCGCGCGTGTACGAGGAGGAGGCGCGTCTCGCTCAGCTCGCGAG 1001
QY 667 CTCAAGGCGTGTCTCAGGAGCAGACTCTGTTCACCTTCTACTCGCGCTCGATTCTCTG 726
Db 1002 CTCAAGGCGTGTCTCAGGAGCAAAACCTGTACCTTCTACTCGCGCTCGATTCTGTTC 1061
QY 727 GGCT-----ATGATGCTCTGAGTCGAGCAGCGGAGATGGGGTGGGTGAGG 777
Db 1062 GGTACGACGCAATGCGCGCGCGCGTGTCTCCGAGGTGGAAGCGCGGTGTAAGG 1121
QY 778 GTGAAGCTGGTGGACTTGTGCCATGTGGCGCGAGGTGATGGGTGATTGACCACAACTTC 837
Db 1122 GTGAAGCTGGTGGACTTGTGGCGATGTGACGATGGGAGCGGGTGTATTGACCACAACTTC 1181
QY 838 CTGGCGCGGCTGTGTGCTGTGATCAAGTGTCTGTGACATGTTCCGAGACTCTCTCAT 897
Db 1182 TTGGCGCGGCTGTGTGCTGTGATCAAGTGTCTGTGACATGTTCCGAGGTACCAGG 1241
QY 898 ACGCAGCTTGTGGTCTCTTTAAGA 923
Db 1242 AAGCGCTCTTCAGATCAATCTTTGAAA 1267

RESULT 9
AX513580
LOCUS AX513580 643 bp DNA linear PAT 05-OCT-2002
DEFINITION Sequence 17 from Patent WO02059324.
ACCESSION AX513580
VERSION AX513580.1 GI:23559681
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1
REFERENCE
AUTHORS Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.
TITLE Novel inositol polyphosphate kinase genes and uses thereof
JOURNAL Patent: WO 02059324-A 17 01-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source Location/Qualifiers
1. .643
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
ORIGIN
Query Match 35.0%; Score 470.4; DB 6; Length 643;
Best Local Similarity 95.9%; Pred. No. 1.9e-61;
Matches 509; Conservative 0; Mismatches 16; Indels 4; Gaps 3;

QY 42 TCCCATACCATGTCGACCTCCACCCGCGGAGCACCAGTCGCGCCACCGCGCTC 101
Db 107 TCCCATACCATGTCGACCTCCACCCGCGGAGCACCAGTCGCGCTC 166
QY 102 CGCAGCAGCTGGGCGCGCTCATCGAGGTCGCGCTTCTTCAAGCGGCTCCAGGC 161
Db 167 CGCAGCAGCTGGGCGCGCTCATCGAGGTCGCGCTTCTTCAAGCGGCTCCAGGC 226
QY 162 CGGCGACCGTGGGAGCAGAGTGCCTTCTATGAGGGTCTCTCGCCACCGCGCGT 221
Db 227 CGGCGACCGTGGGAGCAGAGTGCCTTCTATGAGGGTCTCTCGCCACCGCGCGT 286
QY 222 CCGGCGCGGATCCGAGACACCTTCTTCCCGGTTTCCAGCGCAGCGACTCTCCCGAC 281
Db 287 CCGGCGCGGATCCGAGACACCTTCTTCCCGGTTTCCAGCGCAGCGACTCTCCCGAC 346

```

```

QY 282 CGAGCGCAGCCCGCGGAGCCGCACTCTCACTCTCGTCTCGACGACCTCTCTCGCGGGTT 341
Db 347 CGAGCGCAGCCCGCGGAGCCGCACTCTCGCACTCTCTCGTCTCGACGACCTCTCTCGCGGGTT 406
QY 342 TCAGCGCCCTCGCTCGCAGACATCAAGATCGGCGCATACAGTGGCCACCGAGTTCCGC 401
Db 407 TGAGCGCCCTCGCTCGCAGACATCAAGATCGGCGCATACAGTGGCCACCGAGTTCCGC 466
QY 402 CGAGCCCTACATCGCCCAAGTGCCTCGCCAGACGCGCGGACCAAGAGCTTCTCTCGG 461
Db 467 GGAGCCCTACATCGNCAAGTACCTINGCAAGACCGCGGACCAAGAGCTTCTCTCGG 526
QY 462 ATTTCGCGCTCTCGCGCTCGTCCAGTCTCGCGCCCGAGCGCGCTGTGGCGGACCGAGG 521
Db 527 ATTTCGCGCTCT--TGCCTCGAGTCTCTCGCGCCCGAGCGCGCTGTGGCGGACCGAGG 584
QY 522 -CCCGAGGTCAAGGCCATGGACA-CCGCGCGCGTCTCGCGCGCTCTCGG 570
Db 595 CCCCGGGGTGAANGCTATGACACCCGCTCGGNGNCCGCGGCTCTTCGG 635

RESULT 10
AX513581
LOCUS AX513581 519 bp DNA linear PAT 05-OCT-2002
DEFINITION Sequence 18 from Patent WO02059324.
ACCESSION AX513581
VERSION AX513581.1 GI:23559682
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1
REFERENCE
AUTHORS Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.
TITLE Novel inositol polyphosphate kinase genes and uses thereof
JOURNAL Patent: WO 02059324-A 18 01-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source Location/Qualifiers
1. .519
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
ORIGIN
Query Match 22.2%; Score 298.4; DB 6; Length 519;
Best Local Similarity 89.4%; Pred. No. 1.6e-35;
Matches 328; Conservative 0; Mismatches 38; Indels 1; Gaps 1;

QY 627 GTACGGAGAAAAGGTGGAGTCTTGTACAGCTCGCGAGCTCAAGGCGTGGTTCGAGA 686
Db 2 GTACGGAGAAAAGGTGGAGTCTTGTACAGCTCGCGAGCTCAAGGCGTGGTTCGAGG 61
QY 687 GCAGACTCTGTCCACTTCTACTCGCGTCCGATCTTCTCGGGCTATGATGCTGTGAGT 746
Db 62 GCAGACTCTGTCCACTTCTACTCGCGTCCGATCTTCTCGGGCTATGATGCTGTGAGT 121
QY 747 CGCAGCAGCGGAGATGGGGTGGGTGAAGTGGTGGACTTTTGCCTATGTGGC 806
Db 122 CGCAGCAGCGGAGTGGGGTGGGTGAAGTGAAGTGGTGGACTTTTGCCTATGTGGC 181
QY 807 CGAGGTGTAGGGGTGATTGACCACTTCTCGGGCGGCTCTGCTCGCTGATCAAGTT 866
Db 182 CGAGGTGTAGGGGTGATTGACCACTTCTCGGGCGGCTCTGCTGCTGATCAAGTT 241
QY 867 CGTTCTGACATGTTCTCGGAGACTCTCTACACAGCTTTGGTCTCTCTTAAGAGAG 926
Db 242 TGTTCTGACATGTTCTCAGAGACTCTCTCAGACGAGCTTTGGTCTCTCTTAAGAAA 301
QY 927 GATCCTGSCA--TTTCGATTTGATAACAAAGCCCTTACAGTTTCTTGTGAAAAAGAGCG 985
Db 302 GATCCTGSCATTTTCGATTTGATAACAAAGGAAACATCTTTCAGCTGCCAAAAAANCA 361

```



```
QY 986 CCTCCGA 992
Db 362 CCAGTGA 368

RESULT 11
AX513582
LOCUS AX513582 353 bp DNA linear PAT 05-OCT-2002
DEFINITION Sequence 19 from Patent WO02059324.
ACCESSION AX513582
VERSION AX513582.1 GI:23559683
KEYWORDS
SOURCE
ORGANISM
Eucalyptus grandis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1
REFERENCE
AUTHORS Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.
TITLE Novel inositol polyphosphate kinase genes and uses thereof
JOURNAL Patent: WO 02059324-A 19 01-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
Location/Qualifiers
1..353
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN
Query Match 14.3%; Score 191.6; DB 6; Length 353;
Best Local Similarity 90.2%; Pred. No. 2.1e-19;
Matches 238; Conservative 0; Mismatches 21; Indels 5; Gaps 3;
QY 667 CTCAGGCGTGGTTCGAGGAGCAGACTCTGTTCACACTCTACTCGCGCTCGATTCTTCTG 726
Db 1 CTCAGGCGATGGTTCGAGGAGCAGACTCTGTTCACACTCTACTCGCGCTCGATTCTTCTG 60
QY 727 GCCTATGATGCTGTGAGTGCAGCAGCGGAGATGGGGTGGGGTGCAGGTGAAGCTG 786
Db 61 GCCTATGATGCTGTGAGTGCAGCAGCGGAGATGGGGTGGGGTGAAGCTG 120
QY 787 GTGACATTTGCCATGTGCGCGAGGGTGTATGGGTTCGATTCGAGATTCCTCATACG-C 844
Db 121 GTGACATTTGCCATGTGCGCGAGGGTGTATGGGTTCGATTCGAGATTCCTCATACG-C 180
QY 845 GCCTCTGCTGCTGATCAAGTTC--GTTTCTGACATTTCTCGGAGATTCCTCATACG-C 901
Db 181 AGCTCTGCTAGTCAAGTTCGTTTCTTGACATTTCTCGAGATTCCTCATACG-C 901
QY 902 AGCCTTTGGGTCCTTTTAAAGAGA 925
Db 241 AGCCTTTGGGTCCTTTTAAAGAGA 264

RESULT 12
AX513574
LOCUS AX513574 1195 bp DNA linear PAT 05-OCT-2002
DEFINITION Sequence 11 from Patent WO02059324.
ACCESSION AX513574
VERSION AX513574.1 GI:23559674
KEYWORDS
SOURCE
ORGANISM
Eucalyptus grandis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots;
rosids; Myrtales; Myrtaceae; Eucalyptus.
1
REFERENCE
AUTHORS Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.
TITLE Novel inositol polyphosphate kinase genes and uses thereof
JOURNAL Patent: WO 02059324-A 11 01-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
Location/Qualifiers
1..1195
/organism="Eucalyptus grandis"
/mol_type="unassigned DNA"
/db_xref="taxon:71139"
116--1048
/organism="Eucalyptus grandis"
/mol_type="unassigned DNA"
/db_xref="taxon:71139"
116--1048
/organism="unassigned protein product"
/codon_start=1
/protein_id="CAD52956.1"
/db_xref="GI:23559675"
/db_xref="RENTREMBL:CAD52956"
/translation="MLKVPDQVAGHGRDGGKLGPLVDDSGRFYKPLQSDHGRDTEVA
EYESFYNTPEPHIRKEFPAPHGTKTIEASDGGSPQPHIVLELIVSRTPSLMDIK
TGSRTWPEASEEYIQKLEKDRNSTSVSLGRVYONLSBAGFWQPEKKVYISF
NADGVSALRKPFVSSNLSPNVDPCLLYASKYVYCHRGGLIHLQLKEWFEVQNYH
FYCSLILIDRESALDGCAPKVLVDFAHVMGHDGVIDHNFLLGLGSLVKIFIRIDIA
DEDNKCAKCEVNLKENGFKYSTPELDEAC"
ORIGIN
Query Match 13.7%; Score 184.6; DB 6; Length 1195;
Best Local Similarity 55.6%; Pred. No. 2e-18;
Matches 471; Conservative 0; Mismatches 349; Indels 27; Gaps 5;
QY 56 CCGACCTCCACCCCGCGAGCACCAAGTCGCGGCGCCCGCGCTCCGCCAGCAAGCTGG 115
Db 114 CCATGCTCAAGTCCCGATCATCAAGTCGCGGTACCGGGGAGACGGGGAAGCTGG 173
QY 116 GCCCGCTCATCGACGGCTCCGGGCTCTTTCTACAAAGCGCTCCAGCGCGGACCGCTGGG 175
Db 174 GGCCACTGGTGGATGATTGGGGCGCTTCTATAAGCCTCTCCAGAGCGATCATCGCGAG 233
QY 176 AGCAGAGGTGCGCTTCTATGAGCGCTTCTCCGCCACGCGCGCTCCGCCCGCATCC 235
Db 234 ACACGGAAGTGGCGCTTTTACGAGTCACTTATTCATACCGAGATCCAGGTCACTTC 293
QY 236 GAGACACCTTCTCCCGCGGTTCACGCGACGCGACTCTCCCCACCGAGCGGAGCCCG 295
Db 294 G---AAATCTTCTCGCTTCGCTTACGGAATAAGACTAT---TGAGCGCTCTGATGAT 347
QY 296 GGGAGCGCATCTTCACCTCGCTCGACGACCTCTCGGGGGGTTTCAGCGCGCTTCGCG 355
Db 348 CGGCTCTCAACCTCACCTGCTTCGAGGATCTCGTCTCGGTCGCGACGAAACCACTC 407
QY 356 TCGCAGACATCAAGATCGGGCCATCACTGGCCACCGAGTTCGCGGAGCCCTACATCG 415
Db 408 TCATGGACATCAAGACTGGATCCAGAAATGATTCGCGAGGCTCTGAGGAGTACATPC 467
QY 416 CCAAGTGCCTCGCCAAAGACCGCGGACCGAGCGCTTCTCGTCTCGATTCGCGCTCTCG 475
Db 468 AAAAGTCTTAGAGAAAGATCGAAATAGCACACGCTTCTATGGGTTTATGAGATTTCTG 527
QY 476 GCCTCCGAGTGTCTCG---CCCCGAGGGCGCGTGTGCGGACGAGCGCGCGAGGTGA 532
Db 528 GGCTAAGGGTATATCAAAATAGCGAAGCTGGATTTTGGCAACCTGAGAAAGGTTGTT 587
QY 533 AGGCCATGGACACCGCGGGTTCGCGCGCTTCGCGCGCTACGTGTCTATCC----- 585
Db 588 ATAGCTTTAATTCGCGAGCGGTTCAGGTTCGGTCTGAGGAAAGTTTGTCTTCTCAACTGT 647
QY 586 ----GTTTCGCGAGAGGGATGACATGTGGCTCGCGCGCGCGGTGTACCGAGGAAAG 640
Db 648 CTCGGGGTCAAAATGTGATCCGGATTTGTTGTATGATCAAAAGTTTACTGTACCGGG 707
QY 641 GTGAGTCTTTGTACAGCTCGCGAGCTCAAGCGGTGGTTCGAGGAGCAGACTCTGTTC 700
Db 708 GTGGAATTTGGCAACAATGCTTCAGCTGAAGAAATGTTTGAAGTTTTCAGACGAATATC 767
QY 701 ACTTCTACTCGCGCTCGATTCTTCTGGGCTATGATGCTGTGCTGAGTTCGACAGCGGAG 760
Db 768 ACTTCTATTTCTGTCTACTCATTTATATGACAGGAGTCTGCTTTGGACGCGC----- 823
QY 761 ATGGGGGTGGGGTTCACCGGTGAAGCTGTGTGCTATTTGCCCATGTGGCGAGGTTGATGGG 820
Db 824 --TGTCGACACCCGAAAGTTAACTGGTGGACTTTGACATGTGATGATGATGCCACGGCG 881
```

```
QY      821  TGATTGACCAACTTCCTGGCGGCTCTGCTGCTGATCAAGTCTGTTCTGACATTG 880
Db      892  TGATGATCAACAATTCCTGGTGGCCCTCTGTTCTGTAATCAAGTTATACGTGACATTG 941
QY      881  TTCGGGA 887
Db      942  CTGATGA 948

RESULT 13
LUS310150      25054 bp      DNA      linear      PLN 24-JAN-2002
Locus          Linum usitatissimum variety Bombay Ngc-D, Ngc-A and Ngc-B genes.
DEFINITION     AJ310150
ACCESSION      AJ310150
VERSION        AJ310150.1 GI:13509206
KEYWORDS       Ngc-A gene; Ngc-A protein; Ngc-B gene; Ngc-B protein; Ngc-D gene;
              Ngc-D protein.
SOURCE         Linum usitatissimum (flax)
ORGANISM       Linum usitatissimum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Malpighiales; Linaceae; Linum.
REFERENCE      1
AUTHORS        Dodds,P.N., Lawrence,G.J. and Ellis,J.G.
TITLE          Contrasting modes of evolution acting on the complex N locus for
              rust resistance in flax
JOURNAL        Plant J. 27 (5), 439-453 (2001)
MEDLINE        21461288
PUBMED        11576428
REFERENCE      2 (bases 1 to 25054)
AUTHORS        Dodds,P.N.
TITLE          Direct Submission
JOURNAL        Submitted (22-MAR-2001) Dodds P.N., Plant Industry, CSIRO, GPO Box
              1600, Canberra, ACT 2601, Australia
FEATURES       Location/Qualifiers
               1..25054
               /organism="Linum usitatissimum"
               /mol_type="genomic DNA"
               /variety="Bombay"
               /db_xref="taxon:4006"
               /haplotype="N"
               join(2881..3431,3533..4652,4770..5045,5140..6519)
               /gene="Ngc-D"
               join(2881..3431,3533..4652,4770..5045,5140..6519)
               /gene="Ngc-D"
               /note="N rust resistance gene: TIR-NBS-LRR protein"
               /codon_start=1
               /product="Ngc-D protein"
               /protein_id="CAC35321.1"
               /db_xref="GI:13509207"
               /db_xref="GOA:Q9ARCA"
               /db_xref="SPTREMBL:Q9ARCA"
               /translation="MMRSDSGSDISSEHSVSDPTLLPLPSGEVEVFLSPRGDVRK
TFADHLYTSLVRSKIRTRDEEGQKGETIGPSLIQAITESKIYIPILPTQNYASSKWC
LQELAKVNCWKGGEAKGQHIIPVYFMDPRDVRHPDPSGPKYKESFEQHLKDDPET
LLEWKALQEVGMKGWHISLTGQGAVDKIFTEVELHLRANTYLTATDELVDGDSV
DEMYKLLNDSTSEKIIGYMGRLGTLTATAVYKVSQMFERRCCFLDRIETLLKN
DGVALQNKVTSIDILRKDFCOAKNADGVQVIRVSRHKIFVVDVWNSFRDDIF
GKLTAFSADSRFLVTRDARTLELRGCKLFKHEGMSHSLKLFKSHAFGVDPED
YASICEFVQVSGSLPLAKVIGSLLPTEKSPKDKLJELKAPVNVQYRLKISYN
ELTDNEKQILDVACLVEFGAKKEIPITWSDCGFYPTTIRTLVQSLVIRINDEEF
MHDRIQDGRAIVEKQYKRSIRIWSNDAILKRNKENDCCEALRVMRGEFA
LTNDEEFGTQNRLEVLNGDLGNFNVLPSLEWLVRYHGDGCPGSLNKLMLLEL
EVSVDTSWEGNEIKAAKGLKVHLMCKGLEKVPDLSTCRLELLRFSICRMHGE
LDIENKDLVLDIFOIRITALKEVESLQNLQDQVSSGLIEVPAGISKLSSLYL
LNINIKDKVTLPNGLKILLISFSLSALPSSILFRLDVRYSTNLRPLNLSVTNLT
RLRUEVGHIGIPGLGELKECLFLDAENLNDGLENLVLKELAVECRILEKL
PSALAEKTKHKLVTGQCNILGIEYGLANLGSLSHLETSGCPLTVVESHSLNLTGT
LELSGYTITLPPSLITYKLKSLKVSQSDSPDLTNKRLCKIKCGCNFIITGL
HTLSLEELRMVMSIRKLDLTGLVLEILQFDSCTQLTETIRGIGLESQRLHMSRC
QSIKELNLSGLKILSYILLEKCRHLKEVYRLBELKELDFNTHRG"
              17600..18457
               /function="unknown"
               /note="ORF
              related to human inositol hexakisphosphate kinase 2
              protein"
               /codon_start=1
               /product="hypothetical protein"
               /protein_id="CAC35324.1"
               /db_xref="GI:13509210"

variation
```

```
/db_xref="GOA:Q9ARC1"
/db_xref="SPTREMBL:Q9ARC1"
translation="MLKPEHOVAGHODINGLLPLVDDSGRFYKPLQDDEGATEMA
FTYSTTKNTIPQXVRRFQAFQGTQSIASDGLHPLHLLDITSLRELHLCVMDIK
IGSRWYEPASAYIEKCLKUDVESNELLPGRISGLQYVNGKSESSEVILKPERKLL
LNLNADEVLLKRVFNSKSDQPCFAAVVSSNGILAQLLELKAWEKDFQTIYH
FNSCVLMLEYEKETVMNGLSGESLGAALKLIDFAHVTGNGVVDHFNGLGLCSLIK
FISENQIML"
join(20265..20815,21104..22202,22415..22690,22818..24218)
/gene="Ngc-B"
join(20265..20815,21104..22202,22415..22690,22818..24218)
/gene="Ngc-B"
/note="N rust resistance gene homolog: TIR-NBS-LRR
protein"
/codon_start=1
/product="Ngc-B protein"
/protein_id="CAC3325.1"
/db_xref="GI:13509211"
/db_xref="GOA:Q9AR40"
/db_xref="SPTREMBL:Q9AR40"
translation="MWRSDTSSIGSFSLRSVDPPTLLPSGEVEVFLSPGPDVRO
TFADHLVAVLRSKIRTRDEREGLQGETIGSSLIQALTESKIVIPILTONYASKWC
LOELAKMVDCKWNGGKGQHLILPVFPMPDRVHRDPSPYKEAFQHNKLDPEI
LLEWEALQDVGMKGWHNELTGGQAVVDKIFTTIEFHLRANYTLATDELVLGDISSV
EVEMLMDHDSERIIIGYMGGLGKTLAKAVFNKVSQFOFECDDNIRETLR
NDGVVALQNVISDILRQSDQAKNASDGVRIIRVRHKIFVVDIDIDESFHFDEI
FGKLDGSDSRFLITRDARTLELLNECKMFGLEMGHSDHSLQLFKSHAFGVDPPE
DYASICEEFIVASGLPALKVIGSLFKDKRPWEDKILKAIKPSAKVQERLKVSY
NELTHNEKOIFLDIACLVGAKEVPMWMSDCLYPASTLRTLVORSLVRMDNKKF
WMDHDIRLGRITIVRENSQNPYKRSRIWSNDAIDILKNREGNDVVALRVDMMKGG
YALINKEFKFSRLRFLVNLGDSGNFNILPNLRWVRVYRGDPSPGLNKLNLVIL
ELDGYVTHSWKGNNEIKAAGLKVNLTSQILEKVPDLSTCRGLELLCFHKQMMR
GELDITGFKDLKVDINTEITTLKGVESLQNLQDLVGRSGLEIEVAGISKLSSLE
FLDITSVKHDEVMPLNGKLVILSSFSLSLPSLKLIDICDNRLOLNLASVTN
LTRLHLKVGHEIPGLKGLKLESLSICNAPNLNDLGNLVLKELALERCPIILG
KLPSLAETLKLHVIRWCDVLGEIYHGNLSDLSHLDISWCPRLTVMDDLHLLKL
GTLVSSGELTNLPLSUSITKTLKLEVRSSQDPLNLKNLRDLITITGRELIEHA
GLHTLESLEELSMERCPVRKLDLAGLKXTHICTHICQTLTEIRGLGSLQMLEF
SGCOSIKELPNLGNLKLKFLSKBEKRLQKEVNGLELEWLDFTDRRLKLKYLKLV
SRKQGKLVTSQARRALFINN"
ORIGIN

Query Match 13.5%; Score 181.6; DB 8; Length 25054;
Best Local Similarity 55.4%; Pred. No. 3.6e-18;
Matches 474; Conservative 0; Mismatches 349; Indels 33; Gaps 5;

QY 50 CCAUTGCGACCTCCACCGCGGAGACCAAGTCGCGCGCACCGCGCTCGCGCACA 109
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 110 AGCTGGGCGCGTCAATGACGCGCTCGCGCTCTTCTTACAAGCGCTCCAGCGCGCGGACC 169
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 170 GTGGGAGACAGAGTGCCTTCTATGAGCGTTCCTCGCGCACCGCGCGCTCGCGGCC 229
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 230 GCATCCGAGACACCTTCCTCCCGGTTCCAGGACGCGACTCTCTCCCGACCGAGCGC 289
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 290 AGCCGGGAGCGCATCTCTCACTCTGCTCTGACGACCTCTCTCGCGGGTTTCAGGCGC 349
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 350 CTGGGTCGACAGATCAAGATCGCGGCATCACTGAGCGCATCACTGAGCGCTCGGTCGATC 9898
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 9899 CATGTGTATGACATCAAGATTTGGTTCCAGGACATGTGTATCCCGAGGCTTCTCAAGCCT 9958
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 410 ACATCGCAAGTGCTCGCCAGGACCGCGGACACAGCGGTCTGTCTCGGATCCGCG 469
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 9959 ACATCGAAGATGCTGAGAGGATGTGGAATCAAGCAGTCCCTTCTCGGGTTTAGGA 10018
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
QY 470 TCTCGCGGTCGAGTCTGTCGSCCCCGA-----GGCGCGGTGTGCGGACGG 517
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10019 TATCCGGATTGAGGTTTACGGTAACGACAAAGAAATCATCTGAGTGGTCTTTGAAGCCTG 10078
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 518 AGCGCCGCGAGGTGAAGCCCATGACACCGCGCGGTCCGCGCGTGTCTCCGCGGTACG 577
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10079 AGAGGAAGCTTCTCAGAAATCTTACTGCTGATGAAGTTAGGCTGGTCTTAAACGTTTG 10138
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 578 TGTCAATCGTTCGCGA---CGAGGGGATGAGTGTGCGCTCGCCGCGCGGTGTACGAG 634
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10139 TTTCTCCAAACCGGAAGTCGACCAACCGGATTTGTTCTGTTGACGGGTGCTTTACGGT 10198
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 635 GAAAAGGTGGAGTCTTGTACAGCTGCGGAGCTCAAGCGGTGTTCCGAGGACGAGATC 694
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10199 GTTCAATGGGATCTTGGCTCAGTTCTGAGCTGAAGCGGTGTTGAGGATCAGACCA 10258
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 695 TGTTCACCTTCTACTCGGCTCGATTTCTTCTGGGCTATGATGCTGCT-----G 742
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10259 TCTATCAATTCACCTTCTTCTTCTTCTTCTATGCTATGAGAAGGAGAACGAGATGG 10318
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 743 CAGTCGACGACGCGGAGATGGGGTGGGTGACGCTGAAGCTGTGACTTGTCCCATG 802
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10319 TGAATGGAGGAAGAGAGATTTGGGTGTTGCTGCGGTGAAGCTCATCGATTTGCTCATG 10378
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 803 TGGCGAGGCTGATGGGTGATGACCAACTTCTCGCGCGGCTCTGCTCGCTGATCA 862
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10379 TGACGAGGAATGGCGTTATCGATCACAATTTCTTGGAGGGCTCTGCTCATTGATAA 10438
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 863 AGTTCGTTCTGACAT 878
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10439 AGTTCATTTCCGAGAT 10454
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
AX513572 1105 bp DNA linear PAT 05-OCT-2002
LOCUS AX513572
DEFINITION Sequence 9 from Patent WO02059324.
ACCESSION AX513572
VERSION AX513572.1 GI:23559672
KEYWORDS Glycine max (soybean)
SOURCE Glycine max
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE 1
AUTHORS Shi,J., Beach,L.R., Wang,H., Rafalski,J.A. and Cahoon,R.E.
TITLE Novel inositol polyphosphate kinase genes and uses thereof
JOURNAL Patent: WO 02059324-A 9 01-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source
1..1105
/organism="Glycine max"
/mol_type="unassigned DNA"
/db_xref="taxon:3847"
12..851
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD52955.1"
/db_xref="GI:23559673"
/db_xref="REMBL:CAD52955"
translation="MLKIPHOVAGHKADGILGLVDDFGKFKYKPLQTKDDTRGS
TELSFYSLAAAHDSIRSFPFAPHGTRILLDSDGSPHPHIVLELLCGYKPSVM
DWKIGTWHLGDSEYICKLKADRESSLPIGRISGVKDSISSSEPKRSIQCLIS
HAGVALVNLKFSVNSNHDHDPACAFATEYVAGVLRLQKLDKEVQTVYFYS
SLVWYVEKDLGKATNPLKLVDFAHVVDGNGVIDHFNGLGSLCFIKFLKDLILAV
LHK"
ORIGIN

Query Match 13.3%; Score 179; DB 6; Length 1105;
Best Local Similarity 55.6%; Pred. No. 1.4e-17;
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2004, 23:13:14 ; Search time 601 Seconds
(without alignments)
9500.130 Million cell updates/sec

Title: US-10-042-894A-7
Perfect score: 1344
Sequence: 1 gcacgagtcagtcgctcac.....ataaaaaaaaaaaaaaaaaa 1344

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 29Jan04:*
1: geneseqn1980s:*
2: geneseqn1980s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1344	100.0	1344	6	AAD43514
2	864.4	64.3	922	6	AAD43513
3	844.8	62.9	1169	6	AAD43511
4	837.4	62.3	923	6	AAD43512
5	740.4	55.1	3416	6	AAD43522
6	686	51.0	899	6	AAD43518
7	470.4	35.0	643	6	AAD43519
8	298.4	22.2	519	6	AAD43520
9	191.6	14.3	353	6	AAD43521
10	184.6	13.7	1195	6	AAD43516
11	179	13.3	1105	6	AAD43515
12	149.8	11.1	1020	6	AAD43517
13	148.6	11.1	1168	3	AAC33023
14	147	10.9	1104	3	AAC33685
15	147	10.9	1243	3	AAC38692
16	145.6	10.8	1130	3	AAC48750
17	142	10.6	1130	3	AAC40313
18	104.6	7.8	464	6	ABL93575
19	72.4	5.4	113193	7	AAD54645
20	71.4	5.3	135638	7	ABX34289
21	68.6	5.1	88421	6	AAL40781
22	65.4	4.9	484	9	ADE82058
23	65	4.8	897	7	ADA71279

ALIGNMENTS

RESULT 1
AAD43514
ID AAD43514 standard; DNA; 1344 BP.
XX
AC AAD43514;
XX
DT 14-NOV-2002 (first entry)
XX
DE Maize inositol polyphosphate kinase (IPPK) DNA #4.
XX
KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
KW nutritional value; animal feed; transgenic; gene; ds.
XX
OS Glycine max.
XX
FH Key
FT CDS
FT 52..921
FT /*tag= a
FT /product= "Maize IPPK protein #4"
XX
XX WO200259324-A2.
XX
XX 01-AUG-2002.
XX
XX 09-JAN-2002; 2002WO-US003120.
XX
XX 12-JAN-2001; 2001US-0261465P.
XX
XX (FION-) PIONEER HI-BRED INT INC.
XX
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX WPI; 2002-636540/68.
XX P-PSDB; AAE26196.
XX
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
XX useful in modulating phytic acid biosynthesis by decreasing phytate or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed.
XX
XX Claim 1; Page 64-65; 86pp; English.
XX
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
XX polypeptides and polynucleotides. Sequences of the invention are useful
XX in modulating the phytic acid biosynthesis by decreasing phytate and/or
XX increasing non-phytate phosphorous to improve the nutritional value of

C	24	63.8	4.7	5452	9	ADC86736	Adc86736 Human GPC
	25	63.4	4.7	3849	4	AAF25795	Aaf25795 S. chryso
	26	63.4	4.7	9975	7	AAL61173	Aal61173 Actinosyn
C	27	63.4	4.7	82746	7	ABL61224	AbL61224 Actinosyn
	28	63.2	4.7	1227	6	ABL61294	AbL61294 N. unifor
	29	63.2	4.7	24379	2	AAT93095	Aat93095 Streptomy
	30	63.2	4.7	24379	2	AAV25925	Aav25925 Streptomy
	31	62.2	4.6	1084	9	ADB78935	AdB78935 Rice tran
	32	62	4.6	2000	7	ADA71938	Ada71938 Rice gene
	33	62	4.6	5877	6	ABS78681	Abs78681 Kitasatos
	34	62	4.6	27541	4	AAD17185	Aad17185 Streptomy
	35	62	4.6	125401	4	AAD17186	Aad17186 Streptomy
	36	61.2	4.6	3354	7	AAL61181	Aal61181 Actinosyn
	37	61	4.5	53789	2	AAV21187	Aav21187 Amycolato
	38	61	4.5	114955	2	AAx53491	Aax53491 Human ade
	39	60.2	4.5	1266	7	AAD53058	Aad53058 Streptomy
	40	59.8	4.4	1313	4	AAF74867	Aaf74867 Leishmani
C	41	59.6	4.4	114955	2	AAx53491	Aax53491 Human ade
	42	59.4	4.4	1047	7	ABZ66716	Abz66716 Orthosomy
	43	59.4	4.4	45055	7	ABZ66808	Abz66808 Orthosomy
	44	59.4	4.4	65140	4	AAD17184	Aad17184 Streptomy
C	45	59.2	4.4	1000	3	AAA02484	Aaa02484 Human col

CC animal feed, or to reduce the environmental impact of animal waste.
CC Polynucleotides of the invention are to produce transgenic plants with an
CC altered phenotype. IPPK proteins are used to screen compounds that
CC modulate their activity and raising anti-idiotypic antibodies. The
CC present sequence is maize IPPK DNA
XX
SQ Sequence 1344 BP; 228 A; 426 C; 399 G; 291 T; 0 U; 0 Other;
Query Match 100.0%; Score 1344; DB 6; Length 1344;
Best Local Similarity 100.0%; Pred. No. 1.7e-271;
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCACGAGGTCAGTCCGTGACCCCTCGCGCCCATAGTCCCTTCCCATACCATGTCGAC 60
Db 1 GCACGAGGTCAGTCCGTGACCCCTCGCGCCCATAGTCCCTTCCCATACCATGTCGAC 60
QY 61 CTCACCCCGCGAGACCAAGTCGCGCCGACCGCGCTCCGCCAGCAGCTGGGCCG 120
Db 61 CTCACCCCGCGAGACCAAGTCGCGCCGACCGCGCTCCGCCAGCAGCTGGGCCG 120
QY 121 CTCATCGAGGTCGCGGCTCTTCTACAGCCGCTCCAGCCGCGACCGTGGGAGCAC 180
Db 121 CTCATCGAGGTCGCGGCTCTTCTACAGCCGCTCCAGCCGCGACCGTGGGAGCAC 180
QY 181 GAGGTGCGCTTCTATGAGCGGTTCTCCGCCACGCGCGCTCCGCCCGCATCCGAGAC 240
Db 181 GAGGTGCGCTTCTATGAGCGGTTCTCCGCCACGCGCGCTCCGCCCGCATCCGAGAC 240
QY 241 ACCTTCTTCCCGGTTCCACGCGACCGGACTCTCTCCACGAGCGCGAGCCCGGGAG 300
Db 241 ACCTTCTTCCCGGTTCCACGCGACCGGACTCTCTCCACGAGCGCGAGCCCGGGAG 300
QY 301 CCGCATCTCTCACTGCTGCTCGACGACCTCTCGCGGGGTTTCAGGCGCCCTGCGTCCGA 360
Db 301 CCGCATCTCTCACTGCTGCTCGACGACCTCTCGCGGGGTTTCAGGCGCCCTGCGTCCGA 360
QY 361 GACATCAAGATCGGCGCCATCAGTGGCCACCGAGTTCGCGGAGCCCTACATCGCCAG 420
Db 361 GACATCAAGATCGGCGCCATCAGTGGCCACCGAGTTCGCGGAGCCCTACATCGCCAG 420
QY 421 TGCCTCGCAAGAACCGCGGACACGAGCTTCTGCTCGGATTCGCGCTCCGCGCTC 480
Db 421 TGCCTCGCAAGAACCGCGGACACGAGCTTCTGCTCGGATTCGCGCTCCGCGCTC 480
QY 481 CGAGTCGTGCGGCCCGAGCGCGCTGTCGCGACCGAGCGCCCGAGGTGAAGCCCATG 540
Db 481 CGAGTCGTGCGGCCCGAGCGCGCTGTCGCGACCGAGCGCCCGAGGTGAAGCCCATG 540
QY 541 GACACCGCGCGGTCGCGCGCTGCTCCGCGCTACGTGTCTATCCGTTCCGACGAGGGG 600
Db 541 GACACCGCGCGGTCGCGCGCTGCTCCGCGCTACGTGTCTATCCGTTCCGACGAGGGG 600
QY 601 ATGGAGTGTGCGTCCCGCGCGGTTGACGAGGAAAGTGGAGTCTGTGTCACAGCTG 660
Db 601 ATGGAGTGTGCGTCCCGCGCGGTTGACGAGGAAAGTGGAGTCTGTGTCACAGCTG 660
QY 661 CCGAGCTCAAGCGGTGTCGAGGAGCAGACTCTGTTCCACTTCTACTCGCGCTGATT 720
Db 661 CCGAGCTCAAGCGGTGTCGAGGAGCAGACTCTGTTCCACTTCTACTCGCGCTGATT 720
QY 721 CTTCTGGGCTATGATGCTGTCAGTCGACGAGCGAGATGGGGTGGGTGACGGTG 780
Db 721 CTTCTGGGCTATGATGCTGTCAGTCGACGAGCGAGATGGGGTGGGTGACGGTG 780
QY 781 AGCTGGTGAATTTGCCCATGTCGAGGAGGTGATGAGGATGATGACCAACTTCCTG 840
Db 781 AGCTGGTGAATTTGCCCATGTCGAGGAGGTGATGAGGATGATGACCAACTTCCTG 840
QY 841 GCGGGCTCTGCTCGTGTACGATGCTGTTCTGACATGTTCCGAGACTCTCATACG 900
Db 841 GCGGGCTCTGCTCGTGTACGATGCTGTTCTGACATGTTCCGAGACTCTCATACG 900
QY 901 CAGCCTTTGGGTCCTTCTTAAGAGAGGATCTCGGCATTTTCGATTTGATAACAAAGCCCTA 960

Db 901 CAGCCTTTGGGTCCTTCTTAAGAGAGGATCTCGGCATTTGATTTGATAACAAAGCCCTA 960
QY 961 CAAGTTTGTCTGGAAGAAAGCGCTCCGAGTTGTCTGGGTGTGAGATCTGAGACG 1020
Db 961 CAAGTTTGTCTGGAAGAAAGCGCTCCGAGTTGTCTGGGTGTGAGATCTGAGACG 1020
QY 1021 GTCGTCGCGCCACTGTTGCTGCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCT 1080
Db 1021 GTCGTCGCGCCACTGTTGCTGCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCT 1080
QY 1081 CTTTTTTTCGCAACCCCTTACTTCCGAAGAAACTTTTTTTTCCCACTTTGGGGGTGGA 1140
Db 1081 CTTTTTTTCGCAACCCCTTACTTCCGAAGAAACTTTTTTTTCCCACTTTGGGGGTGGA 1140
QY 1141 TTAOCTTGATGCTGTTTGTGCACTCGGTGAGAGTTGTAAGCATGAGGAGGCGGTGT 1200
Db 1141 TTAOCTTGATGCTGTTTGTGCACTCGGTGAGAGTTGTAAGCATGAGGAGGCGGTGT 1200
QY 1201 TGATCCGCAACTGTCAGTCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGC 1260
Db 1201 TGATCCGCAACTGTCAGTCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGC 1260
QY 1261 TGGGATCCGATGTGCTGAGATCGTAGTGTGATGATGATGATGATGATGATGATGATG 1320
Db 1261 TGGGATCCGATGTGCTGAGATCGTAGTGTGATGATGATGATGATGATGATGATGATG 1320
QY 1321 TCCGATAAAAAAATAAAAAA 1344
Db 1321 TCCGATAAAAAAATAAAAAA 1344
RESULT 2
AAD43513
ID AAD43513 standard; DNA; 922 BP.
XX
AC AAD43513;
DT 14-NOV-2002 (first entry)
XX
DE Maize inositol polyphosphate kinase (IPPK) DNA #3.
XX
KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
KW nutritional value; animal feed; transgenic; gene; ds.
XX
OS Zea mays.
FH Key Location/Qualifiers
FT CDS 53..922
FT /*tag= a
FT /product= "Maize IPPK protein #3"
XX
WO200259324-A2.
XX
PD 01-AUG-2002.
XX
PF 09-JAN-2002; 2002WO-US003120.
XX
PR 12-JAN-2001; 2001US-0261465P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX
DR WPI; 2002-636540/68.
DR P-PSDB; AA26195.
XX
PT New inositol polyphosphate kinase polynucleotides and polypeptides,
PT useful in modulating phytic acid biosynthesis by decreasing phytate or
PT increasing non-phytate phosphorous to improve the nutritional value of
PT animal feed.
XX
PS Claim 1; Page 62-63; 86pp; English.

XX The invention relates to novel inositol polyphosphate kinase (IPPK)
CC polypeptides and polynucleotides. Sequences of the invention are useful
CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or
CC increasing non-phytate phosphorous to improve the nutritional value of
CC animal feed, or to reduce the environmental impact of animal waste.
CC Polynucleotides of the invention are used to produce transgenic plants with an
CC altered phenotype. IPPK proteins are used to screen compounds that
CC modulate their activity and raising anti-idiotypic antibodies. The
XX present sequence is maize IPPK DNA
SQ Sequence 922 BP; 137 A; 327 C; 288 G; 170 T; 0 U; 0 Other;

Query Match 64.3%; Score 864.4; DB 6; Length 922;
Best Local Similarity 98.6%; Pred. No. 3e-171;
Matches 869; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 42 TCCCATACCATGTCGACCTCCACCGCGGAGACCAAGTCGCGGACCGCGCCTC 101
DB TCCCATACCATGTCGACCTCCACCGCGGAGACCAAGTCGCGGTCACCGCGCCTC 102
QY 102 CCGCAGCAGCTGGGCGCGCTCATCGAAGCTCCGCGCTTCTTACAGCGCTCCAGGC 161
DB CCGCAGCAGCTGGGCGCGCTCATCGAAGCTCCGCGCTTCTTACAGCGCTCCAGGC 162
QY 162 CGCGACCGTGGGAGCAGAGTGCCTTCTATGAGCGTTCGCGCCACCGCGCGT 221
DB CGCGACCGTGGGAGCAGAGTGCCTTCTATGAGCGTTCGCGCCACCGCGCGT 222
QY 222 CCGCGCGCGATCCGAGACACCTTCTCCCGGTTTCAACGCGACGCGACTCTCCCGAC 281
DB CCGCGCGCGATCCGAGACACCTTCTCCCGGTTTCAACGCGACGCGACTCTCCCGAC 282
QY 282 CGAGCGCAGCCCGGGAGCCGATCCCTCACTCGTCTTCTCGACGACCTCTCGCGGTT 341
DB CGAGCGCAGCCCGGGAGCCGATCCGACCTCGTCTTCTCGACGACCTCTCGCGGTT 342
QY 342 TCAGCGCGCTCGTTCGAGACATCAGATCGCGCGCATCAGTGGCCACCGAGTTCGC 401
DB TCAGCGCGCTCGTTCGAGACATCAGATCGCGCGCATCAGTGGCCACCGAGTTCGC 402
QY 402 GGAGCGCTTACATCGCAAGTGCCTCGCAAGACCGCGGACACGAGCGTTCCTCGG 461
DB GGAGCGCTTACATCGCAAGTGCCTCGCAAGACCGCGGACACGAGCGTTCCTCGG 462
QY 462 ATTCGCGTCTCCGCGTCCGAGTCTCGCGCCCGAGGCGCGGTGTGCGGACCGGAGCG 521
DB ATTCGCGTCTCCGCGTCCGAGTCTCGCGCCCGAGGCGCGGTGTGCGGACCGGAGCG 522
QY 522 CCGGAGTGAAGGCATGGACACCGCGCGTTCGCGCGCGTGTCTCGCGCGTACGTTC 581
DB CCGGAGTGAAGGCATGGACACCGCGCGTTCGCGCGCGTGTCTCGCGCGTACGTTC 582
QY 582 ATCCGTTCCGACGAGGAGTGGATGTGCGTTCGCGCGCGGCGGTGTACGAGGAAAGG 641
DB ATCCGTTCCGACGAGGAGTGGATGTGCGTTCGCGCGCGGCGGTGTACGAGGAAAGG 642
QY 642 TGGAGTCTTGTACAGCTGCGGAGCTCAAGCGGTGTGCGAGGAGCAGACTCTGTCCA 701
DB TGGAGTCTTGTACAGCTGCGGAGCTCAAGCGGTGTGCGAGGAGCAGACTCTGTCCA 702
QY 702 CTCTACTCGCGGTGATCTTCTGGCTATGATCTCTGAGTTCGAGCAGCGGAGG 761
DB CTCTACTCGCGGTGATCTTCTGGCTATGATCTCTGAGTTCGAGCAGCGGAGG 762
QY 762 TGGGGTGGGTGACGCTGAAGCTGTGACTTTTCCCATGTGCGCGGAGGTGATGGGT 821
DB TGGGGTGGGTGACGCTGAAGCTGTGACTTTTCCCATGTGCGCGGAGGTGATGGGT 822
QY 822 GATTGACCAACAATTCTTCTGGCGGGCTCTGCTCGCTGATCAAGTTCGTTCTGACATGT 881
DB GATTGACCAACAATTCTTCTGGCGGGCTCTGCTCGCTGATCAAGTTCGTTCTGACATGT 882

QY 882 TCCGAGAGCTCTCTATACGAGCCTTTGGTCTCTTTAA 921
DB 883 TCCAGAGCTCTCTACAGCGAGCCTTTGGTCTCTTTAA 922

RESULT 3
AAD43511
ID AAD43511 standard; DNA; 1169 BP.
AC AAD43511;
XX 14-NOV-2002 (first entry)
XX Maize inositol polyphosphate kinase (IPPK) DNA #1.
XX Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
KW nutritional value; animal feed; transgenic; gene; ds.
XX Zea mays.
OS
FH Key Location/Qualifiers
FT CDS 84..806
FT /*tag= a
FT /product= "Maize IPPK protein #1"
XX WO200259324-A2.
XX 01-AUG-2002.
XX 09-JAN-2002; 2002WO-US003120.
XX 12-JAN-2001; 2001US-0261465P.
XX (PION-) PIONEER HI-BRED INT INC.
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX WPI; 2002-636540/68.
XX P-PSDB; AAE26193.
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
PT useful in modulating phytic acid biosynthesis by decreasing phytate or
PT increasing non-phytate phosphorous to improve the nutritional value of
PT animal feed.
XX Claim 1; Page 58-59; 86pp; English.
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
CC polypeptides and polynucleotides. Sequences of the invention are useful
CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or
CC increasing non-phytate phosphorous to improve the nutritional value of
CC animal feed, or to reduce the environmental impact of animal waste.
CC Polynucleotides of the invention are used to produce transgenic plants with an
CC altered phenotype. IPPK proteins are used to screen compounds that
CC modulate their activity and raising anti-idiotypic antibodies. The
XX present sequence is maize IPPK DNA
SQ Sequence 1169 BP; 238 A; 362 C; 347 G; 222 T; 0 U; 0 Other;

Query Match 62.9%; Score 844.8; DB 6; Length 1169;
Best Local Similarity 97.2%; Pred. No. 3.9e-167;
Matches 890; Conservative 0; Mismatches 22; Indels 4; Gaps 3;

QY 42 TCCCATACCATGTCGACCTCCACCGCGGAGACCAAGTCGCGGACCGCGCCTC 101
DB TCCCATACCATGTCGACCTCCACCGCGGAGACCAAGTCGCGGTCACCGCGCCTC 133
QY 102 CGCAGCAAGCTGGCGCGCTCATCGACGCTCCGGCTTCTTACAGCGCTCCAGGC 161
DB CGCAGCAAGCTGGCGCGCTCATCGACGCTCCGGCTTCTTACAGCGCTCCAGGC 193
QY 162 CGGCGACCGTGGGAGCAGAGTTCGCTTCTATGAGCGTTCGCGCCACCGCGCGT 221

Db 523 CCGGAGGTGAAGGCTATGACACCGTCGCGCTCGCGCGCTACGTGTC 582
Qy 582 ATCCG-TTCCGACGAGGGATGGAAGTGTGCGTTCGCCGCGGTTACGAGGAAG 640
Db 583 ATCCGTTTCCGACGAGGGATGGAAGTGTGCGTTCGCCGCGGTTACGAGGAAG 642
Qy 641 GTGAGTCTTGTTCACAGCTGCGCGAGCTCAAGCGGTGTTTCGAGGAGCAGACTGTCTCC 700
Db 643 GTGAGTCTTGTTCACAGCTGCGCGAGCTCAAGCGGTGTTTCGAGGAGCAGCTGTCTCC 702
Qy 701 ACTTACTACGCGGTGATTTCTTGGGTATGATGCTGTGAGTCGAGTCGAGAGCGGAG 760
Db 703 ACTTACTACGCGGTGATTTCTTGGGTATGATGCTGTGAGTCGAGTCGAGAGCGGAG 762
Qy 761 ATGGGGGTGGGTGACGCTGAAGCTGGTGGACTTTGCCCATGTGCGCGAGGTGATGGG 820
Db 763 GTGGGGGTGGGTAAACAGTGAAGCTGGTGGACTTTGCCCATGTGCGCGAGGTGATGGG 822
Qy 821 TGATTGACACCAACTTCTTGGGCGGCTCTGCTGCTGATCAAGTTTCTGATG 880
Db 823 TGATTGACACCAACTTCTTGGGCGGCTCTGCTGATCAAGTTTCTGATG 882
Qy 881 TTCGGAGACTCTCATACGAGCCTTTGGGTCTCTTAA 921
Db 883 TTCCAGAGACTCTCATACGAGCCTTTGGGTCTCTTAA 923

RESULT 5

AAD43522
ID AAD43522 standard; DNA; 3416 BP.
XX
AC AAD43522;
XX
DT 14-NOV-2002 (first entry)
XX
DE Maize inositol polyphosphate kinase (IPPK) DNA #9.
XX
KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
KW nutritional value; animal feed; transgenic; gene; ds.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT 72..407
CDS /*tag= a
FT /product= "Maize IPPK protein #6"
XX
FN WO200259324-A2.
XX
PD 01-AUG-2002.
XX
PF 09-JAN-2002; 2002WO-US003120.
XX
PR 12-JAN-2001; 2001US-0261465P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX WPI; 2002-636540/68.
DR P-PSDB; AAE26201.
XX
PT New inositol polyphosphate kinase polynucleotides and polypeptides,
PT useful in modulating phytic acid biosynthesis by decreasing phytate or
PT increasing non-phytate phosphorous to improve the nutritional value of
PT animal feed.
XX
PS Claim 1; Page 75-77; 86pp; English.
XX
CC The invention relates to novel inositol polyphosphate kinase (IPPK)
CC polypeptides and polynucleotides. Sequences of the invention are useful
CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or

CC increasing non-phytate phosphorous to improve the nutritional value of
CC animal feed, or to reduce the environmental impact of animal waste.
CC Polynucleotides of the invention are to produce transgenic plants with an
CC altered phenotype. IPPK proteins are used to screen compounds that
CC modulate their activity and raising anti-idiotypic antibodies. The
XX present sequence is maize IPPK DNA
SQ Sequence 3416 BP; 893 A; 830 C; 856 G; 837 T; 0 U; 0 Other;
Query Match 55.1%; Score 740.4; DB 6; Length 3416;
Best Local Similarity 90.2%; Pred. No. 3.1e-145;
Matches 862; Conservative 0; Mismatches 31; Indels 63; Gaps 4;
Qy 2 CACGAGGTGAGTCCGTCACCCCTCGCGCCCATAGTCCCTTCCCATACCATGTCGAC 61
Db 28 CACCATCGCCACCCCTGTCACCCCTTCTCCATAG-----TCCCATACCATGTCGAC 81
Qy 62 TCCACCCCGCGAGACCAAGTCGCGGCGCACCGCGCTTCCGCCAGCAAGCTGGGCCCGC 121
Db 82 TCCACCCCGCGAGACCAAGTCGCGGCTACCGGCTTCCGCCAGCAAGCTGGGCCCGC 141
Qy 122 TCATCGACGCTCGGCTCTTTTACAAAGCCGCTCCAGCCCGGCGAACGTCGGGAGACG 181
Db 142 TCATCGACGACTCTGGCTCTTCTACAAAGCCGCTCCAGCCCGGCGAACGTCGGGAGC 201
Qy 182 AGGTGCGCTTATGAGGCTTCTCCGCCAGCGCGCTCCCGCCCGCATCCGAGACA 241
Db 202 AGGTGCGCTTATGAGGCTTCTCCGCCAGCGCGCTCCCGCCCGCATCCGAGACA 261
Qy 242 CTTCTTCCCGGTTCCACGCGACGACTCTCTCCACCGAGCGCAGCCCGCGGAGC 301
Db 262 CTTCTTCCCGGTTCCACGCGACGACTCTCTCCACCGAGCGCAGCCCGCGGAGC 321
Qy 302 CGCATCTCACTCTCTCGACGACTCTCTCGGGGTTTGAAGCGGCTTGGTGGGAGC 361
Db 322 CGCATCTCACTCTCTCGACGACTCTCTCGGGGTTTGAAGCGGCTTGGTGGGAGC 381
Qy 362 ACATCAAGATCGGCTTCCAGCGACTCTCTCCACCGAGCGCAGCCCGCGGAGC 421
Db 382 ACATCAAGATCGGCTTCCAGCGACTCTCTCCACCGAGCGCAGCCCGCGGAGC 405
Qy 422 GCCTCGCCACGAGACCGCGGACACGAGCGTCTCTCGGATTCGCGCTCTCCGCGTCC 481
Db 406 -----GACACGAGCGTCTCTCGGATTCGCGCTCTCCGCGTCC 446
Qy 482 GAGTCGTGCGCCCGAGCGCGCTGTGCGGACGAGCGCCCGAGGTGAAGCCATGG 541
Db 447 GAGTCGTGCGCCCGAGCGCGCTGTGCGGACGAGCGCCCGAGGTGAAGCCATGG 506
Qy 542 ACACCGCGCGCTCCGCGCGTCTCCGCGCTACGTGTATCCG-TTGCAGCAGAGGG 600
Db 507 ACATTGTGCGCTCCGCGCGTCTCCGCGCTACGTGTATCCGCTTCCGAGCGGG 566
Qy 601 ATGAGCTGTGCGCTCCGCGCGTGTACGAGGAAAGGTGAGTCTTCTCAGCTG 660
Db 567 ATGAGCTGTGCGCTCCGCGCGTGTACGAGGAAAGGTGAGTCTTCTCAGCTG 626
Qy 661 CGCAGCTCAAGGCTGGTTCGAGGAGCAGACTCTTCTCAGCTCTCTCTCTCGCGT 720
Db 627 CGCAGCTCAAGGCTGGTTCGAGGAGCAGACTCTTCTCAGCTCTCTCTCTCGCGT 686
Qy 721 CTTCTGGCTATGATGCTGCTGAGTCGAGCGAGGATGGGGTGGGTCAGCGTG 780
Db 687 CTTCTGGCTATGATGCTGCTGAGTCGAGTCGAGCGAGGATGGGGTGGGTAACATG 746
Qy 781 AAGCTGTGAGCTTTGCCCATGTGCGCGAGGTCATGGGGTGAATTGACCAACTTCCT 840
Db 747 AAGCTGTGAGCTTTGCCCATGTGCGCGAGGTCATGGGGTGAATTGACCAACTTCCT 806
Qy 841 GCGCGGCTCTGCTGCTGATCAAGTTCGTTCTGACATTTGTCGGAGACTCTCTCAAC 900
Db 807 GCGCGGCTCTGCTGCTGATCAAGTTCGTTCTGACATTTGTCGGAGACTCTCTCAAC 866


```
XX 09-JAN-2002; 2002WO-US003120.
PF
XX
XX 12-JAN-2001; 2001US-0261465P.
PR
XX
XX (PION-) PIONEER HI-BRED INT INC.
PA
XX
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
PI
XX WPI; 2002-636540/68.
DR
XX
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
PT useful in modulating phytic acid biosynthesis by decreasing phytate or
PT increasing non-phytate phosphorous to improve the nutritional value of
PT animal feed.
XX
XX Claim 1; Page 74; 86pp; English.
PS
XX
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
CC polypeptides and polynucleotides. Sequences of the invention are useful
CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or
CC increasing non-phytate phosphorous to improve the nutritional value of
CC animal feed, or to reduce the environmental impact of animal waste.
CC Polynucleotides of the invention are to produce transgenic plants with an
CC altered phenotype. IPPK proteins are used to screen compounds that
CC modulate their activity and raising anti-idiotypic antibodies. The
CC present sequence is maize IPPK DNA
XX
XX Sequence 643 BP; 92 A; 261 C; 178 G; 103 T; 0 U; 9 Other;
SQ
Query Match 35.0%; Score 470.4; DB 6; Length 643;
Best Local Similarity 95.9%; Pred. No. 6.8e-89;
Matches 509; Conservative 0; Mismatches 18; Indels 4; Gaps 3;
QY 42 TCCCATACCATGTCGACCTCCAGCCCGCGGAGACCAAGTCGCCGCGCACCGCGCTC 101
Db 107 TCCCATACCATGTCGACCTCCAGCCCGCGGAGACCAAGTCGCCGCGCACCGCGCTC 166
QY 102 CGCCAGCAAGTGGCGCCGCTCATCGACGGTCCGGGCTCTTCTACAGCGCTCCAGGC 161
Db 167 CGCCAGCAAGTGGCGCCGCTCATCGACGGTCCGGGCTCTTCTACAGCGCTCCAGGC 226
QY 162 CGCGGACGCTGGGAGACGACGAGTGGCTTCTATGAGGGTTCGCCGCCACGCGCGGT 221
Db 227 CGCGGACGCTGGGAGACGACGAGTGGCTTCTATGAGGGTTCGCCGCCACGCGCGGT 286
QY 222 CCGGCGCGCATCGAGACACCTTCTTCCCGGTTCCAGGACGCGACTCTCTCCCGAC 281
Db 287 CCGGCGCGCATCGAGACACCTTCTTCCCGGTTCCAGGACGCGACTCTCTCCCGAC 346
QY 282 CGAGGCGACGCCGGGAGCGCATCTCTACCTCTCTCGACACCTCTCTCGCGGGGT 341
Db 347 CGAGGCGACGCCGGGAGCGCATCGCACCTCTCTCTCGACACCTCTCTCGCGGGGT 406
QY 342 TCAGGCGCGCTCGTTCGACACATCAAGATCGGGGCCATCAGTGGGCCACCGAGTTGCC 401
Db 407 TGAAGCGCCCTGGCTCGACACATCAAGATCGGGGCCATCAGTGGGCCACCGAGTTGCC 466
QY 402 GGAGCCCTACATCGCCAAAGTGCCTCGCCAGGACCGCGGGACCAAGAGCGTCTGTCGG 461
Db 467 GGAGCCCTACATCGNCAAGTACCTTNGCCAAAGGACCGCGGGACCAAGAGCGTCTGTCGG 526
QY 462 ATTCGCGCTCTCGGCGTTCGAGTCGTTCGGCCCGAGGGGCCGCTGTGGGCGGACGAGCG 521
Db 527 ATTCGCGCTCT - TGGCTTCGAGTCGTTCGGCCCGAGGGGCCGCTGTGGGCGGACGAGCG 584
QY 522 -CCGAGGTCGAAGGCCATGACA - CGCGCGCGCTTCGCCGCGGTCTCGG 570
Db 585 CCGCGGGGTGAANGCTATGACACCGGTTCGGNGCCGGGNGTGTCTCGG 635
```

RESULT 8
AAD43520

```
ID AAD43520 standard; DNA; 519 BP.
AC AAD43520;
XX
DT 14-NOV-2002 (first entry)
XX
DE Maize inositol polyphosphate kinase (IPPK) DNA #7.
XX
KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
KW nutritional value; animal feed; transgenic; ds.
XX
OS Zea mays.
XX
FN WO200259324-A2.
XX
XX 01-AUG-2002.
XX
XX 09-JAN-2002; 2002WO-US003120.
XX
XX 12-JAN-2001; 2001US-0261465P.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX WPI; 2002-636540/68.
XX
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
PT useful in modulating phytic acid biosynthesis by decreasing phytate or
PT increasing non-phytate phosphorous to improve the nutritional value of
PT animal feed.
XX
XX Claim 1; Page 75; 86pp; English.
XX
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
CC polypeptides and polynucleotides. Sequences of the invention are useful
CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or
CC increasing non-phytate phosphorous to improve the nutritional value of
CC animal feed, or to reduce the environmental impact of animal waste.
CC Polynucleotides of the invention are to produce transgenic plants with an
CC altered phenotype. IPPK proteins are used to screen compounds that
CC modulate their activity and raising anti-idiotypic antibodies. The
CC present sequence is maize IPPK DNA
XX
XX Sequence 519 BP; 144 A; 109 C; 126 G; 125 T; 0 U; 15 Other;
```

```
Query Match 22.2%; Score 298.4; DB 6; Length 519;
Best Local Similarity 89.4%; Pred. No. 6.1e-53;
Matches 328; Conservative 0; Mismatches 38; Indels 1; Gaps 1;
QY 527 GTACGGAGGAAAAGGTGGAGTCTTGTACAGCTGCGGAGCTCAAGCGGTGTTTCGAGGA 686
Db 2 GTACGGAGGAAAANGTGGAGTCTTGTACAGCTGCGGAGCTCAANGCGTGTTCGAGGG 61
QY 687 GCAGACTCTGTTCCACTTCTACTCGGCGTGCATCTTCTGGGTATGATCTGCTGCAGT 746
Db 62 GCAGACTCTGTTCCACTTCTACTCGGCGTGCATCTTCTGGGTATGATCTGCTGCAGT 121
QY 747 CGCAGCAGGCGGAGATGGGGGTGCGGTGACGGTGAAGCTGGTGACTTTGCCCATGTGGC 806
Db 122 CGCAGCAGGCGGANGTGGGGGTGCGGTGAACAGTGAAGCTGGTGACTTTGCCCATGTGGC 181
QY 807 CGAGGCTGATGGGGTGAATGACCAAACTTCTTGGGCGGGCTCTGCTCGCTGATCAAGTT 866
Db 182 CGAGGCTGATGGGGTGAATGACCAAACTTCTTGGGCGGGCTCTGCTTANCTGATCAAGTT 241
QY 867 CGTTTCTGACATTTGTTCCGAGACTCTCTCATACGACGCTTTGGGTCTTCTTTAAGAGAG 926
Db 242 TGTTTCTGACATTTGTTCCAGAGACTCTCTCAGACGCGAGCTTTGGGTCTTCTTTAAGAAA 301
QY 927 GATCCTGGCA - TTTGATTTGATAA CAAGCCCTCA CAAGTTTGTCTGGAAGAACGCG 985
Db 302 GATCCTGGCATTTTCGATTTGATAA CAAGAACGAGCACTTTTCAGCTGCCAAAAAANCA 361
```

```
QY 986 CCTCCGA 992
Db 362 CCAGTGA 368

RESULT 9
AAD43521
ID AAD43521 standard; DNA; 353 BP.
XX
AC AAD43521;
XX
DT 14-NOV-2002 (first entry)
XX
DE Maize inositol polyphosphate kinase (IPPK) DNA #8.
XX
KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
KW nutritional value; animal feed; transgenic; ds.
XX
OS Zea mays.
XX
PN WO200259324-A2.
XX
PD 01-AUG-2002.
XX
PF 09-JAN-2002; 2002WO-US003120.
XX
PR 12-JAN-2001; 2001US-0261465P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX
DR WPI; 2002-636540/68.
XX
PS Claim 1; Page 75; 86pp; English.
XX
CC The invention relates to novel inositol polyphosphate kinase (IPPK)
CC polypeptides and polynucleotides. Sequences of the invention are useful
CC in modulating the phytic acid biosynthesis by decreasing phytate and/ or
CC increasing non-phytate phosphorous to improve the nutritional value of
CC animal feed, or to reduce the environmental impact of animal waste.
CC Polynucleotides of the invention are to produce transgenic plants with an
CC altered phenotype. IPPK proteins are used to screen compounds that
CC modulate their activity and raising anti-idiotypic antibodies. The
CC present sequence is maize IPPK DNA
XX
SQ Sequence 353 BP; 81 A; 78 C; 91 G; 96 T; 0 U; 7 Other;

Query Match 14.3%; Score 191.6; DB 6; Length 353;
Best Local Similarity 90.2%; Pred. No. 1.2e-30;
Matches 238; Conservative 0; Mismatches 21; Indels 5; Gaps 3;

QY 667 CTCAGGCTGTCTGAGGAGCAGACTCTGTTCCACTTCTACTCGCGCTCGATTCTCTG 726
Db 1 CTCAGGCAATGTTGAGGAGCAGACTCTGTTCCACTTCTACTCGCGCTCGATTCTCTG 60
QY 727 GGTATGATGCTGCTGCTGAGTCCAGCGGAGATGGGGTGGGGTGACGGTGAAGCTG 786
Db 61 GGTATGATGCTGCTGCTGAGTCCAGCGGAGTGGGGTGGGGTGAACGTGAAGCTG 120
QY 787 GTGGACTTTGCCATGTGCGCCAGGGTGATGGGT--GATTGACCAACTTCTCTGGCG 844
Db 121 GTGGACTTTGCCATGTGCGCCAGGGTGATGGGT--GATTGACCAACTTCTCTGGCG 180
QY 845 GGTCTGTGCTGATCAATGTC--GTTTCTGACATTTCTCGGAGACTCTCTATAG-C 901
Db 181 AGCTCTGTGCTGATCAATGTC--GTTTCTGACATTTCTCGGAGACTCTCTATAGCGC 240

QY 902 AGCCTTGGTCTCTTCTTAAGAGA 925
Db 241 AGCCTTGGTCTCTTCTTAAAAA 264

RESULT 10
AAD43516
ID AAD43516 standard; DNA; 1195 BP.
XX
AC AAD43516;
XX
DT 14-NOV-2002 (first entry)
XX
DE Eucalyptus grandis inositol polyphosphate kinase (IPPK) DNA.
XX
KW Enzyme; inositol polyphosphate kinase; IPPK; phytic acid; transgenic;
KW nutritional value; animal feed; gene; ds.
XX
OS Eucalyptus grandis.
XX
PH Key Location/Qualifiers
FT CDS 116..1048
FT /*tag= a
FT /product= "E. grandis IPPK protein"
XX
PN WO200259324-A2.
XX
PD 01-AUG-2002.
XX
PF 09-JAN-2002; 2002WO-US003120.
XX
PR 12-JAN-2001; 2001US-0261465P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX
DR WPI; 2002-636540/68.
XX
DR P-F5DB; AAE26198.
XX
PS New inositol polyphosphate kinase polynucleotides and polypeptides,
XX useful in modulating phytic acid biosynthesis by decreasing phytate or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed.
XX
PS Claim 1; Page 68-70; 86pp; English.
XX
CC The invention relates to novel inositol polyphosphate kinase (IPPK)
XX polypeptides and polynucleotides. Sequences of the invention are useful
XX in modulating the phytic acid biosynthesis by decreasing phytate and/ or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed, or to reduce the environmental impact of animal waste.
XX CC Polynucleotides of the invention are to produce transgenic plants with an
XX altered phenotype. IPPK proteins are used to screen compounds that
XX modulate their activity and raising anti-idiotypic antibodies. The
XX present sequence is Eucalyptus grandis IPPK DNA
XX
SQ Sequence 1195 BP; 342 A; 240 C; 294 G; 319 T; 0 U; 0 Other;

Query Match 13.7%; Score 184.6; DB 6; Length 1195;
Best Local Similarity 55.6%; Pred. No. 4.3e-29;
Matches 471; Conservative 0; Mismatches 349; Indels 27; Gaps 5;

QY 56 CCGACTTCACCCCGGAGCACCAGTGGCCGGCCACCGGCTCTCCGCCAGCAGCTGG 115
Db 114 CCATGCTCAGGTCCCGGATCATCAAGTCGCGGTCAACCGGGGAGACGGGGAAGCTGG 173
QY 116 GCCCGCTCATCGAGGCTCCGGCTCTTCTACAGCCGCTCCAGCCGCGCCACGCTGGG 175
Db 174 GCCCAGTGGTGATGATTCGGGCGCGCTCTATAAGCCTCTCCAGAGCATCATCGGAG 233
QY 176 AGCAGAGGTGCGCTTCTATAGGCGTTCCTCCGCCACCGCGCGTCCCGCCCGCATCC 235
```

```
Db 234 ACACGGAAGTGGCTTTTACAGTCAATCTATTCATACCGAGATCCAGGTCAATTC 293
Qy 236 GAGACACCTTCTTCCCGGTTCCACGCGACGACTCTCCACCGAGCGGAGCCG 295
Db 294 G---CAAAATCTTCTCGTTCAGCGAACTAAGACTAT--TGAGCGTCTGATGAT 347
Qy 296 GGGAGCGGCACTCTCACTCTCTCTCGAGGACTCTCGCGGTTTCAGGCGGCTGCG 355
Db 348 CGGTCCTCAACCTCACCTGTTCTGGAGATCTGCTCGGTCGCAAGAACCATCTC 407
Qy 356 TCGCAGACATCAAGATCGCGCCATCAGTGGCCACCGAGTTCGCGGAGCCCTACATCG 415
Db 408 TCATGACATCAAGACTGGATCCAGACATGATATCGGAGCCCTCTGAGAGTACATCC 467
Qy 416 CCAAGTGGCTCCGCAAGAACCGGGACCAAGAGGTTCTCTCGGATTCGCGTTCGCG 475
Db 468 AAAAGTGTCTAGAGAAAGATCGAAATAGCAAGCGTTTCATTTGGGTTTATAGGATTTCTG 527
Qy 476 GCGTCGAGTCTGTCG---CCCGAGGCGCGGTGTGCGGACCGAGCGCCCGAGGTGA 532
Db 528 GGCTAAGGGTATATCAAAATAGCGAAGCTGATTTTGGCAACCTGAGAGAAAGTGTGTT 587
Qy 533 AGGCATGGACACCGCGCGCTCCGCGCGCTCGCTCGGCTACGTGTCTATCC----- 585
Db 588 ATAGCTTTAATGCGGACGCTGAGTCTGAGTCTGAGAAAGTTGTTTCTTCCACTGT 647
Qy 586 -----GTTGCGCAGAGGGATGAGTGTGCTGCTCGCGCGCGGTGTCGAGGAAAG 640
Db 648 CTCGGGTCCAAATGTGATCGGATTTGTTGTATGATCAAAAGTTTACTGTCAACCGGG 707
Qy 641 GTGAGTCTTCTCAGCTGCGGAGCTCAAGGCGTGGTTCGAGGACGAGACTCTGTTC 700
Db 708 GTGGAATTTGGCAATGCTTACGTCGAGGAAGTGTGTTGAGTTCGAGCAATATC 767
Qy 701 ACTTCTACTCGCGTCGATTTCTTGGGCTATGATGCTGTGCTGAGTTCGACAGCGGAG 760
Db 768 ACTTCTATCTTGTCTACTATTAATTAATGACAGGAGTCTGCTTTGACGGC---- 823
Qy 761 ATGGGGTGGGGTGACGCTGAGCTGAGCTGAGGACTTTGCCCATTGTCGCGGAGGTGATGGG 820
Db 824 --TGTCACACCCGAAAGTAAACTGTTGAGTCTTTCACATGATGATGATGATGATGATG 881
Qy 821 TGATGACCAACTCTCTGCGCGGCTCTCTGCTGATCAAGTCTGTTCTGACATG 880
Db 882 TGATGATCACAACTCTCTGCGGCTCTGCTGATCAAGTCTGTTCTGATCAAGTCTGATGATG 941
Qy 881 TTCGGGA 887
Db 942 CTGATGA 948
```

RESULT 11

AAD43515
ID AAD43515 standard; DNA; 1105 BP.

AC AAD43515;

XX 14-NOV-2002 (first entry)

DE Soybean inositol polyphosphate kinase (IPPK) DNA.

XX Soybean; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
KW nutritional value; animal feed; transgenic; gene; ds.

XX Glycine max.

XX Key Location/Qualifiers
PH 12..851
CDS

FT /*tag= a
FT /product= "Soybean IPPK protein"

XX WO200259324-A2.

```
XX 01-AUG-2002.
XX 09-JAN-2002; 2002WO-US003120.
XX 12-JAN-2001; 2001US-0261465P.
XX (PION-) PIONEER HI-BRED INT INC.
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX WPI; 2002-636540/68.
XX P-PSDB; AAE26197.
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
XX useful in modulating phytic acid biosynthesis by decreasing phytate or
XX increasing non-phytate phosphorus to improve the nutritional value of
XX animal feed.
XX Claim 1; Page 66-67; 86pp; English.
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
XX polypeptides and polynucleotides. Sequences of the invention are useful
XX in modulating the phytic acid biosynthesis by decreasing phytate and/or
XX increasing non-phytate phosphorus to improve the nutritional value of
XX animal feed, or to reduce the environmental impact of animal waste.
XX Polynucleotides of the invention are to produce transgenic plants with an
XX altered phenotype. IPPK proteins are used to screen compounds that
XX modulate their activity and raising anti-idiotypic antibodies. The
XX present sequence is soybean IPPK DNA
XX Sequence 1105 BP; 279 A; 311 C; 233 G; 282 T; 0 U; 0 Other;
```

Query Match 13.3%; Score 179; DB 6; Length 1105;

Best Local Similarity 55.6%; Pred. No. 6.3e-28;

Matches 460; Conservative 0; Mismatches 340; Indels 27; Gaps 5;

```
Qy 61 CTCACCGCGGAGCACCAGTCGCGCGGCGCACCGCGCTCCGACAGCTGGGCGG 120
Db 15 CTCAAGATCCCGAGCACCAGGTGGCGGCGCACAGGCCAAGGACGGAATCTTGGGCCA 74
Qy 121 CTCATCAGCGCTCCGCGCTCTTCTACAAGCGCTTCCAGGCGCGGCGGAGGAGC 180
Db 75 CTCGTGACAGATTGGAATTTCTAAGCCCTCCAGACCAACAAGACGAGACACC 134
Qy 181 GAGTCTGCTTCTATGAGGCGTT-----CTCGCCCAACGCGCGCTCCGCGCGGATC 234
Db 135 CGCGCTCCACCGAACTCTCTTTTACACCTCTCTCGCGCGCGCGCGGACGACTACTCC 194
Qy 235 CGAGACACCTTCTTCCCGGTTCCAGCGACGCGACTCTCCCGCACGAGGCGCAGCGC 294
Db 195 ATCGCTCTCTTCTCCCGGCTTTCAGCGACCCGCTCTCTGAGCGCTCCGACGCGTCC 254
Qy 295 GGGAGCGCGATCTCTACCTCTGCTCTCGACGACTCTCTCGCGGCTTTCAGCGCGCTTC 354
Db 255 GG---TCCCGCCCTCACCTGCTCTGAGGACCTCTCTCTCGGCTACTCCAAACCCCTCC 311
Qy 355 GTCGACACATCAGATCGGCGCATCAGTGGCCACCGAGTTCGCGGAGCCCTACATC 414
Db 312 GTCATGAGCGTAAGATCGGCTCCAGAACCTGGGAGACTCCGAGGACTACATC 371
Qy 415 GCCAAGTCTTCGCAAGGACCGCGGACACGAGCGTCTCTGCTCGGATTCGCGGTCTCC 474
Db 372 TGCAGTCTCTGAGAGAGGACAGAGATCTCTTAGTTCCTTGGCTTTCAGATCTCG 431
Qy 475 GGCCTCCGAGTCTGTCGCGCCCGGAGGCGCGGTGTGGGAGCGGAGCGCGGAGGTGAAG 534
Db 432 GGAGTCAAGGACTCTATCTCTCTCTGGAACCTACAGGAATCTCTCCAGTGTCTATCC 491
Qy 535 GCCATGACACCGCGCGGTCCGCGCGGTGCTCGCGGCT---ACGTGTCTATCGGTGCC 591
Db 492 GCCATGCTGTGCACTTGTCTTCTCAAGAGTTCGTTTCTCTAATAATATCAACCATGAT 551
```


PR	27-AUG-1999;	99US-0151066P.	
PR	27-AUG-1999;	99US-0151080P.	
PR	30-AUG-1999;	99US-0151303P.	
PR	31-AUG-1999;	99US-0151438P.	
PR	01-SEP-1999;	99US-0151930P.	
PR	07-SEP-1999;	99US-0152363P.	
PR	10-SEP-1999;	99US-0153707P.	
PR	13-SEP-1999;	99US-0153758P.	
PR	15-SEP-1999;	99US-0154018P.	
PR	16-SEP-1999;	99US-0154039P.	
PR	20-SEP-1999;	99US-0154779P.	
PR	22-SEP-1999;	99US-0155139P.	
PR	23-SEP-1999;	99US-0155486P.	
PR	24-SEP-1999;	99US-0155659P.	
PR	28-SEP-1999;	99US-0156458P.	
PR	29-SEP-1999;	99US-0156596P.	
PR	04-OCT-1999;	99US-0157117P.	
PR	05-OCT-1999;	99US-0157753P.	
PR	06-OCT-1999;	99US-0157865P.	
PR	07-OCT-1999;	99US-0158029P.	
PR	08-OCT-1999;	99US-0158232P.	
PR	12-OCT-1999;	99US-0158369P.	
PR	13-OCT-1999;	99US-0159293P.	
PR	13-OCT-1999;	99US-0159294P.	
PR	13-OCT-1999;	99US-0159295P.	
PR	14-OCT-1999;	99US-0159329P.	
PR	14-OCT-1999;	99US-0159330P.	
PR	14-OCT-1999;	99US-0159331P.	
PR	14-OCT-1999;	99US-0159637P.	
PR	14-OCT-1999;	99US-0159638P.	
PR	14-OCT-1999;	99US-0159584P.	
PR	21-OCT-1999;	99US-0160741P.	
PR	21-OCT-1999;	99US-0160767P.	
PR	21-OCT-1999;	99US-0160768P.	
PR	21-OCT-1999;	99US-0160770P.	
PR	21-OCT-1999;	99US-0160814P.	
PR	21-OCT-1999;	99US-0160815P.	
PR	22-OCT-1999;	99US-0160980P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0160989P.	
PR	22-OCT-1999;	99US-0161404P.	
PR	25-OCT-1999;	99US-0161405P.	
PR	25-OCT-1999;	99US-0161406P.	
PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161992P.	
PR	28-OCT-1999;	99US-0161993P.	
PR	29-OCT-1999;	99US-0162142P.	
Query Match			11.1%; Score 148.6; DB 3; Length 1169;
Best Local Similarity			53.0%; Pred. No. 1.5e-21;
Matches 453; Conservative			0; Mismatches 369; Indels 33; Gaps 5;
QY	49	ACCATGTCGACCTCCACCGCGGAGACACCAAGTCGCGGCGCACCGGCGCTCCGCCAGC 108	
DB	170	ACCAAGATGACAGTCAAGTCCCTGACATCAGGTGAGGACACATTGCTAAACACGGG 229	
QY	109	AGCTGGGCGCCCTCATCGACCGGCTCCGCCCTCTTCTACAGCGCTCCAGCGCGGAC 168	
DB	230	AAGCCTGGTCTCTCGTAGATGACAAAGGTCCGTTCTTCAAGCCACTTCAGGGCGATTCT 289	
QY	169	CGTGGGACACAGAGTCCCTTCTATGAGGCTTCTCGGCCACCGCGCTCCGCCGCC 228	
DB	290	CGTGGTGAATCGAGTAAAGTTCTACGAATCTTCTCTCTCAACACAGAGGTTCCAGAA 349	
QY	229	CGCATCCGAGACACCTTCTTCCCGGTTCCACGCGACGCGACTCTCTCCACCGAGCGG 288	
DB	350	CACATCC---ATAGATATTTCCGGTGTATCAGGCACTCA-----GCATTT 394	
QY	289	CAGCCCGGGAGCCGACATCCTCACTCGTCTCTCGACGACCTCTCTCGCGGGTTTCAGGCG 348	

DB	395	GAAGGTTCTGATGGAGACGCCATGATGGTGTGGAAAAATCTTCTTCAGAAATACTCAAA 454	
QY	349	CCCTGGTCGAGACATCAAGATCGGCGCCATCACTGTCGCCACCGAGTTTCGCGGAGCCC 408	
DB	455	CCATCAGTAATGGATGTTAAGATGGTTCAGAAACATGGTATCTCTGATGCATCTGAAGAA 514	
QY	409	TACATCGCCCAAGTCTCGCCCAAGACCGCGGAGCACAGGCTTCTGCTCGATTCGCGC 468	
DB	515	TACATCCAAAAATGTTGAAGAAAGACACGGGTACCAACCGTGTCACTCGGGTTTCAGG 574	
QY	469	GTCTCGGCGTCCGAGTCTGTCGCGCCCGAGGGCGCC---GTGTGGCGGACGAGCGCCCG 525	
DB	575	ATCTCTGTTTCGAAGTGTATGATCACAAAGATCGAGTTTCTGGAAGCCCGAGGGAAG 634	
QY	526	GAGGTGAAGGCCATGGACACCGCCGCGTCCGCCGCTGCTCCGGCGCTACGTGCA--- 582	
DB	635	CTTCTTCGCGGCTCGATGTAGATGAGCGAGATTGACTCTGAGGAAGTTTGTATCATCT 694	
QY	583	-----TCGGTTGCCGACGAGGGGATGGACTCTGCGCTCGCGCGCGGCTGATCGGA 633	
DB	695	AACTCACTTTCGACACTGGCTCGAAACCTGACTCTGCTTTGCTTCGAGTGTTCACGGC 754	
QY	634	GGAAAAGGTGGAGTCTTGTCAAGCTGCGCGAGCTCAAGGCGTGGTTCGAGGAGCAGACT 693	
DB	755	GGTTCGCCACGGATCTTAACGAGTTGCTGGAACCTCAAGACCTGGTTCGAGAACCAAG 814	
QY	694	CTGTTCCACTTCTACTCGGCGTCTGATCTTCTGGGCTATGAGCTGCTGCAATC---GGA 750	
DB	815	CTCTACCATTTCAACTCTTGTTCCGATTTTAATGGTCTATGAGATGAATCCATCTTGAAG 874	
QY	751	GCAGCGGAGATGGGGTGGGCTGAGCGGTGAGCTGAGCTGCTGCGCTCGCGCGGCTGACGA 810	
DB	875	GGAAATGATGATGATGCTAGACCAACAGTCAAGCTGGTGGATTTTGCCTCATGTTCTTGAT 934	
QY	811	GGTGATGGGTCATTTGACCAACAACTTCTGCGCGGGCTCTGCTCGCTCATCAAGTTCGTT 870	
DB	935	GGTAATGTTGTCATTTGACCACTTCTTGGGTGGTCTTGTCTCTTTTATAAACTTCATT 994	
QY	871	TCTGACATGTTCCG 895	
DB	995	CGTGAGATTTCTCAG 1009	
RESULT 14			
AAC33685			
ID	AAC33685 standard; DNA; 1104 BP.		
XX	AAC33685;		
XX	17-OCT-2000 (first entry)		
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 3949.		
KW	Hybridisation assay; Genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway; metabolic pathway;		
XX	promoter; termination sequence; ss.		
OS	Arabidopsis thaliana.		
PN	EP1033405-A2.		
PD	06-SEP-2000.		
PF	25-FEB-2000; 2000EP-00301439.		
PR	25-FEB-1999; 99US-0121825P.		
PR	05-MAR-1999; 99US-0123180P.		
PR	09-MAR-1999; 99US-0123548P.		
PR	23-MAR-1999; 99US-0125788P.		
PR	25-MAR-1999; 99US-0126264P.		
PR	29-MAR-1999; 99US-0126785P.		
PR	01-APR-1999; 99US-0127462P.		
PR	06-APR-1999; 99US-0128234P.		

PR	13-OCT-1999;	99US-0159295P.	
PR	14-OCT-1999;	99US-0159329P.	
PR	14-OCT-1999;	99US-0159330P.	
PR	14-OCT-1999;	99US-0159331P.	
PR	14-OCT-1999;	99US-0159637P.	
PR	14-OCT-1999;	99US-0159638P.	
PR	18-OCT-1999;	99US-0159584P.	
PR	21-OCT-1999;	99US-0160741P.	
PR	21-OCT-1999;	99US-0160767P.	
PR	21-OCT-1999;	99US-0160768P.	
PR	21-OCT-1999;	99US-0160770P.	
PR	21-OCT-1999;	99US-0160814P.	
PR	21-OCT-1999;	99US-0160815P.	
PR	22-OCT-1999;	99US-0160980P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0160989P.	
PR	25-OCT-1999;	99US-0161404P.	
PR	25-OCT-1999;	99US-0161405P.	
PR	25-OCT-1999;	99US-0161406P.	
PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161992P.	
PR	28-OCT-1999;	99US-0161993P.	
PR	29-OCT-1999;	99US-0162142P.	
Query Match			10.9%; Score 147; DB 3; Length 1104;
Best Local Similarity			52.9%; Pred. No. 3.1e-21;
Matches 452; Conservative 0; Mismatches 370; Indels 33; Gaps 5;			
QY	49	ACCATGTCGACCTCCACCCCGGAGCACCAAGTCGCGGCCACCGCGCTCCGCGAGC	108
DB	120	AACAAGATGACGTCAAAGTCCCTGAACATCAGGTTGCGAGCACACATTCGTAAGACGGG	179
QY	109	AAGCTGGCGCCGCTCATCGACGGCTCGCGCTCTTCTACAGCGCTCCAGCGCGCGAC	169
DB	180	AAGCTGCTCTCTCTAGATGACAAGGTCGGTTCCTCAAGCCACTTCAGGGCGATTCT	239
QY	169	GTGGGAGCAGAGGTCGGCTTCTATGAGCGGTTCTCCGCCACCGCGCGCTCCCGGC	228
DB	240	CGTGGTGAATTCGAGGTAAGGTTCTAGAAATCTTCTCTCAACACAGAGGTTCCAGAA	299
QY	229	CGCATCCGACACACTTCTTCCCGGTTCCAGCGACGCGACTCTCCCCACCGAGCG	288
DB	300	CACATCC---ATGATATTTCCCGGTATACGGCACTCAA-----GCAGTT	344
QY	289	CAGCCCGGAGCGCATCTCCTCGTCTCGACACTCTCGCGGGTTTCAGGCG	348
DB	345	GAAGGTTCTGATGAGCAGCCATGATGGTGTTCGAAATCTTTCGAGAAATCTCAAAA	404
QY	349	CCCTGGTCGCACACATCAAGATCGGCGCCATCACTGGCCACCGAGTTTCGCGGAGCC	408
DB	405	CCATCAGTAATGATGTTAAGATGGTTCGAGAACATGGTATCTGATGATCTGAAGAA	464
QY	409	TACATCCGCAAGTGCTCGCCAGGACCGCGGACACGAGCGTTCTGCTCGAATTCGCG	468
DB	465	TACATCCAAATGTTTGAAGAAAGACACGGGTACCAACACCGTGTCATCGGGTTTCAG	524
QY	469	GTCTCGGCGTCCAGTCTCGCCCGCGCGCGCC---GTGGGGGACGAGCGCGCG	525
DB	525	ATCTCTGGTTTGAAGTATGATCAAGAATCGAGTTCTGGAAGCCCGAGAGAA	584
QY	526	GAGGTGAAGCCATGACACCGCGCGCTCCCGCGCTCCCGCGCTACGTGCA---	582
DB	585	CTTCTTCGCGGGTCGATGTAGATGAGCGAGATTGACTCTGAGGAAGTTTGTATCATCT	644
QY	583	-----TCGTTCCGACGAGGGATGGACTGTGCGCTCGCGCGCGGTGTACGGA	633
DB	645	AACCTCACTTCGACACTGGCTCGAAACCTGCTCTGCTTTTGCCTCGAGTGTTCAGCG	704
QY	634	GGAAGAGTGGAGTCTTGTACAGCTGCGCGAGCTCAAGCGGTGTTGAGAGCAGACT	693
Db	705	GGTTCACCGGATCTTAACGAGTTGCTGGAACTCAAGACCTGGTTCGAGAACCAACG	764
QY	694	CTGTTCCACTTACTCGGCGTCGATTCTTCTGGGCTATGATGCTGCTCAGTC---GCA	750
DB	765	CTCTACCAATTCACCTCTGTTTCGATTTAAGGTCATAGAGAAATCCATCTTGAAG	824
QY	751	GCAGCGGAGATGGGGTGGGTGACGGTGAAGCTGGTGGACTTTGCCCATGTGCCGAG	810
DB	825	GGAAATGATGATGATGCTAGACCAACAGTCAAGCTGGTGGATTGCTCATGTTCTGAT	884
QY	811	GGTATGGGTGATTGACCAACAACTTCTCGGCGGCTCTGCTGCTGATCAAGTTCGTT	870
DB	885	GGTAATGCTGTCATTGACCACTAACTTCTGGGTGCTTTGCTTTTCAATAAATTCATT	944
QY	871	TCTGACATGTTCCG	885
DB	945	CGTGAGATTCCTCAG	959
RESULT 15			
AAC38692			
ID	AAC38692 standard; DNA; 1243 BP.		
XX	AC AAC38692;		
XX	17-OCT-2000 (first entry)		
XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 21894.		
XX	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway; metabolic pathway;		
KW	promoter; termination sequence; ss.		
XX	Arabidopsis thaliana.		
OS	Arabidopsis thaliana.		
XX	EP1033405-A2.		
XX	06-SEP-2000.		
XX	25-FEB-2000; 2000EP-00301439.		
PR	25-FEB-1999;	99US-0121825P.	
PR	05-MAR-1999;	99US-0123180P.	
PR	09-MAR-1999;	99US-0123548P.	
PR	23-MAR-1999;	99US-0125788P.	
PR	25-MAR-1999;	99US-0126264P.	
PR	29-MAR-1999;	99US-0126785P.	
PR	01-APR-1999;	99US-0127462P.	
PR	06-APR-1999;	99US-0128234P.	
PR	08-APR-1999;	99US-0128714P.	
PR	19-APR-1999;	99US-0129845P.	
PR	19-APR-1999;	99US-0130077P.	
PR	21-APR-1999;	99US-0130449P.	
PR	23-APR-1999;	99US-0130891P.	
PR	28-APR-1999;	99US-0131449P.	
PR	30-APR-1999;	99US-0132048P.	
PR	30-APR-1999;	99US-0132407P.	
PR	04-MAY-1999;	99US-0132484P.	
PR	05-MAY-1999;	99US-0132485P.	
PR	06-MAY-1999;	99US-0132486P.	
PR	06-MAY-1999;	99US-0132487P.	
PR	07-MAY-1999;	99US-0132863P.	
PR	11-MAY-1999;	99US-0134256P.	
PR	14-MAY-1999;	99US-0134218P.	
PR	14-MAY-1999;	99US-0134219P.	
PR	14-MAY-1999;	99US-0134221P.	
PR	14-MAY-1999;	99US-0134370P.	
PR	18-MAY-1999;	99US-0134768P.	
PR	19-MAY-1999;	99US-0134941P.	
PR	20-MAY-1999;	99US-0135124P.	
PR	21-MAY-1999;	99US-0135353P.	
PR	24-MAY-1999;	99US-0135629P.	

PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match	10.9%;	Score 147;	DB 3;	Length 1243;
Best Local Similarity	52.9%;	Pred. No. 3.2e-21;		
Matches 452;	Conservative	0;	Mismatches 370;	Indels 33; Gaps 5;
QY	49	ACCATGTCGACCTCCACCGCGCGGAGCACCAAGTCGCGCGCCACCGCGCTCCGCCAGC	108	
Db	257	AACAAGATGCAGCTCAAAAGTCCCTGAACATCAGGTTCGAGACACATTCCTAAAGACGGG	316	
QY	109	AAGCTTGGGCGCGCTCATCGACGGCTCGCGCTCTTCTACAGCCGCTCCAGGCGCGGCAC	168	
Db	317	AAGCCCTGTCTCTCGTAGATGACAAGGTCGGTCTTTCGAAGCACATTCACGGCGATTCT	376	
QY	169	CGTGGGAGACGAGGTGCGCTCTATGAGCGTTCCTCCGCCACGCGCGCTCCCGGCC	228	
Db	377	CGTGGTGAATCGAGGTAAAGTTCACGAATCTTCTCCTCAACAACAGAGGTTCCAGAA	436	
QY	229	CGCATCCGAGACACTTCTTCCCGGTTCCAGGACGGGACTCTCCGCCACGAGGCG	288	
Db	437	CACATCC---ATAGATATTTCCCGGTGTATCAAGGCACCTCAA-----GCAGTT	481	
QY	289	CAGCCCGGGAGCGCATCTCTCACTCGTCTCTGACGACTCTCCGGGGTTTCAGGCG	348	
Db	482	GAAGTCTTGATGAGCAGCATTGATGCTGTGGAAATCTCTTGCAATACTCAAAA	541	
QY	349	CCCTGCGTCGAGACATCAAGATCGGCGCCATCACTGTGCCACCGATTCGCCGGAGCC	408	
Db	542	CCATCAGTAATGAGTGTAAAGATGGGTTCCAGAAATGTTATCTCTGATCACTGAGAA	601	
QY	409	TACATGCCAAGTCCTCGCCPAGGACCGCGGACCAAGAGCGTTCTGTGCGATTCCGC	468	
Db	602	TACATCCAAAATGTTTGAAGAAGACAACGGTACCACAAACCGTGTATCGGTTTCAGG	661	
QY	469	GTCTCCGCGCTCCGAGTCGTCGCGCCCGAGGGGCGCC---GTGTGGCGGACGAGCGCCCG	525	
Db	662	ATCTCTGGTTTCGAAGTGATGATCAAAAGATCGAGTTTCTGGAAGCCCGAGAGGAAG	721	
QY	526	GAGTGAAGCCCATGACACCGCGCGCTCCGCGCGTCTCCGGCGTACTGTGTCA---582		
Db	722	CTTCTTCGCGGGCTCGATGTAGATGGAGCGAGATTGACTCTGAGGAAGTTTGATCATCT	781	
QY	583	-----TCCGTTCCGACGAGGGGATGGACTGTGCGCTCGCGCGCGGTGTACGGA	633	
Db	782	AACTCACTTTGCACTGGCTCGAAACCTGACTCTGCTTTTGGCTTCAGTGTTTACGGC	841	
QY	634	GGAAAAGTGGAGTCTTGTTCACAGTTCGCGAGCTCAAGGCGTGTCTGAGGACGAGACT	693	
Db	842	GGTTCGCCAGGATCTTAACGCAGTGTCTGGAACTCAAGACCTGTTTCGAAACCAACG	901	
QY	694	CTGTTCACATTCATTCGCGGTCGATCTTCTGGGCTATGATGCTGCTGCAGTC---GCA	750	
Db	902	CTCTACCAATTCAACTCTTGTTCGATTTTAATGGTCTATGAGAAATGATTCATCTTGAG	961	
QY	751	GCAGCGGAGATGGGGTGGGTGACGTTGAAGCTGTGGACTTTGCCCATTGTGCGCGAG	810	
Db	962	GGAAATGATGATGATGCTAGACCAAGTCAAGCTGGTGGATTTTGCTCATGTTCTTGAT	1021	
QY	811	GGTGTGGGGTGAATGACCACAACTTCTCGCGCGGCTCTGCTCGCTGATCAAGTTGTT	870	
Db	1022	GGTAATGGTGTCAATGACCAATCACTTCGTTGGTGTCTTCTCTTTCATAAATCTCAT	1081	
QY	871	TCAGCATTTGTTCCG	885	
Db	1082	CGTGAATCTTCAG	1096	

Search completed: March 27, 2004, 02:03:51
Job time : 610 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2004, 01:44:37 ; Search time 118 Seconds
(without alignments)
6320.797 Million cell updates/sec

Title: US-10-042-894A-7
Perfect score: 1344
Sequence: 1 gcacgaggtcagtcggtcac.....ataaaaaaaaaaaaaaaaaa 1344

Scoring table: IDENTITY NUC

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Processing: Minimum Match 0%
Maximum Match 100%

Maximum RACON 100%
Listing first 45 summaries

Database :

```
Database :
Issued Patents NA:**
1: /cgn2_6/pdata/2/ina/5A COMB.seq:**
2: /cgn2_6/pdata/2/ina/5B COMB.seq:**
3: /cgn2_6/pdata/2/ina/6A COMB.seq:**
4: /cgn2_6/pdata/2/ina/6B COMB.seq:**
5: /cgn2_6/pdata/2/ina/POTUS COMB.seq:**
6: /cgn2_6/pdata/2/ina/backfiles1.seq:**
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	69.6	5.2	1926	4	US-09-249-585A-4		Sequence 4, Appli
2	69.6	5.2	1931	2	US-09-130-114-2		Sequence 2, Appli
3	58.6	4.4	38506	3	US-09-320-878-19		Sequence 18, Appli
4	58.6	4.4	38506	4	US-09-141-908-1		Sequence 1, Appli
5	58.6	4.4	38506	4	US-09-657-404-19		Sequence 19, Appli
6	58.6	4.4	4411529	3	US-09-103-840A-1		Sequence 1, Appli
7	58	4.3	1082	4	US-09-881-165-4		Sequence 4, Appli
8	57.6	4.3	23673	4	US-09-773-816-1		Sequence 1, Appli
9	57	4.2	11220	3	US-09-105-537-32		Sequence 32, Appli
10	57	4.2	36778	3	US-09-105-537-5		Sequence 5, Appli
11	55.2	4.1	15872	3	US-09-105-537-1		Sequence 1, Appli
12	55.2	4.1	15872	4	US-09-091-609-1		Sequence 1, Appli
13	55.2	4.1	15872	4	US-09-091-609-3		Sequence 3, Appli
14	55.2	4.1	4403765	3	US-09-103-840A-2		Sequence 2, Appli
15	54.8	4.1	1704	1	US-08-528-199-2		Sequence 2, Appli
16	54.8	4.1	1704	1	US-08-528-199-5		Sequence 5, Appli
17	54.8	4.1	4403765	3	US-09-103-840A-2		Sequence 2, Appli
18	54.4	4.0	985	4	US-09-056-556-182		Sequence 2, Appli
19	54.4	4.0	985	4	US-09-072-596-177		Sequence 182, App
20	54.4	4.0	985	4	US-09-072-596-177		Sequence 177, App
21	54.4	4.0	4411529	3	US-09-072-967-182		Sequence 182, App
22	53.8	4.0	1333	4	US-09-103-840A-1		Sequence 1, Appli
23	53.4	4.0	921	4	US-09-372-422A-9		Sequence 9, Appli
24	53.4	4.0	3084	4	US-09-252-991A-6598		Sequence 6598, Ap
25	53.4	4.0	3273	4	US-09-252-991A-6639		Sequence 6639, Ap
26	53.2	4.0	6854	4	US-09-352-991A-6578		Sequence 6578, Ap
27	53.2	4.0	152331	3	US-09-194-905-7		Sequence 7, Appli
					US-09-128-155-16		Sequence 16, Appli

ALIGNMENTS

RESULT 1.

```

US-09-249-585A-4
/ Sequence 4, Application US/09249585A
/ Patent No. 6417002
/ GENERAL INFORMATION:
/ APPLICANT: Horlick, Robert
/ TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SE
/ FILE REFERENCE: 0867/0D905
/ CURRENT APPLICATION NUMBER: US/09/249,585A
/ CURRENT FILING DATE: 1999-02-11
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: Patentin version 3.0
/ SEQ ID NO 4
/ LENGTH: 1926
/ TYPE: DNA
/ ORGANISM: Epstein Barr Virus
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)-(1926)
/ OTHER INFORMATION: template strand of EBNA-1 DNA
US-09-249-585A-4

```

Query Match	5.2%;	Score 69.6;	DB 4;	Length 1926;
Best Local Similarity	45.6%;	Pred. No. 1.1e-06;		
Matches 246;	Conservative 0;	Mismatches 294;	Indels 0;	Gaps 0;
39	CTTCCGCCATACCATGTCCGACTCACCAGCGCGGAGCACCAAGTCGGCGGCACCGCGC	98		
274	CCTCGTCTCGGCCTCCCGCTCTCTGTCTCCCCGTCTGTCTCTCTCCCCTGCTCGT	333		
99	CTCGGCAGGAAGCTGGGCGCGCTCATCGAGGCTCCGGCCTCTTTACAAGCGCTCCA	158		
334	CTCTCTCCCGCTCTCCCGCTCTCCCGTCTCTCCCGTCTCTCTCTCTCCCGTCTCT	393		
159	GGCGGCGAGACGTGGGGAGCAGAGGTGCGCTTCTATGAGCGTCTCCGGCCACGCGCG	218		
394	CGTGCTCTCTCCCGCTCTCTCCCGTCTCTCGTCTCTCTCGTCTCTCTCGTCTCT	453		
219	CGTCCCGGCGCGATCCGAGACACTTTCTCCCGGGTTTCA CGGCAAGGACTCTCTCCC	278		
454	CTTCCCGTCTCTCCCGTCTCTCGTCTCTCTCCCGTCTCTCCCGTCTCTCTCGTCTCT	513		
279	CACGAGGCGGAGCGCGGGAGCGCATCTCTACCTCTGTCTCTGACGACTCTCTGGCGG	338		
514	CTCTCTCCCGTCTCTGTCCTCTCTCCCGTCTCTCGTCTCTCTCTCTCTCTCTCTCTCT	573		
339	GTTTCAGGCGCCTCGCTGCGAGACATCAAGATCGGCGGCATCAGCTGGGCACGAGTTC	398		
574	CCCGTCTCTCCCGTCTCTGTCCTCTCTCCCGTCTCTGTCCTCTCCCGTCTCTCCCGTCTCT	633		

;; PRIOR APPLICATION NUMBER: US 60/179,305
;; PRIOR FILING DATE: 2000-01-31
;; NUMBER OF SEQ ID NOS: 1
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 1
;; LENGTH: 23673
;; TYPE: DNA
;; ORGANISM: Human
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (1)...(23623)
;; OTHER INFORMATION: n = A,T,C or G
US-09-773-816-1

Query Match 4.3%; Score 57.6; DB 4; Length 23673;
Best Local Similarity 49.2%; Pred. No. 0.001;
Matches 181; Conservative 0; Mismatches 184; Indels 3; Gaps 1;

QY 311 ACCTGCTCTCGACGACCTCTCGCGGGTTTCAGGCGCCCTGCTGCGATCGACATCAAGA 370
DB 14307 ACCGCGCAGTGTGCGGCTGTGCGCTGCGGCGCTGCTGCGGATCGTCCGATGC 14366
QY 371 TCGGCGCATACAGTGGCCACCGAGTTTCGCGGAGCCCTACATCGCCAAAGTGCCTCGCCA 430
DB 14367 GCGCCACATCCCGCCCAACCGGGTGTCTGAGACCAAGCATGTGACGTGCGCG 14426
QY 431 AGGACCGGGACCAAGAGTTCGTGCGATTTCGCGTCTCGGCGTCTCGGCGTCCGAGTCTCG 490
DB 14427 CGTTCGCGCGAGCGCGCGGACCTCGACCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 490
QY 491 GCGCCGAGGCGCGTGTGCGGAGCGAGCGCGCGGAGTGAAGCGCATGACACCGCGCG 550
DB 14487 CCACAAAGCGCGCGTGTGCGGCTGTGACCTGCGCGCGCGCGCGCGGAAATCGTGGCG 14546
QY 551 GCGTCCGCGCGTGTGCGGCGTGTGCGGCGTGTGCGGCGTGTGCGGCGTGTGCGGCGTGTG 607
DB 14547 GCGTCCGCGTGTGCGGCGTGTGCGGCGTGTGCGGCGTGTGCGGCGTGTGCGGCGTGTG 607
QY 608 GTGCGTCCGCGCGGCGTGTGCGGCGTGTGCGGCGTGTGCGGCGTGTGCGGCGTGTGCGG 14606
DB 14607 ACTGCTGCGCGCGCGTGTGCGGCGTGTGCGGCGTGTGCGGCGTGTGCGGCGTGTGCGG 667
QY 668 TCAAGCGG 675
DB 14667 TCAGCGCG 14674

RESULT 9

US-09-105-537-32
; Sequence 32, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 11220
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-105-537-32

Query Match 4.2%; Score 57; DB 3; Length 11220;
Best Local Similarity 44.7%; Pred. No. 0.0012;
Matches 222; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

QY 66 CCGCGCGGAGCACCAGTTCG 125

DB 8322 CCGCGCGGAGCGTCTCGTGGTGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8381
QY 126 CGACGGCTCCGCGCTCTTCTACAAAGCGCTCCAGGCGCGGAGCGCGCGCGCGCGCGCGCGCG 185
DB 8382 CGAGGCCCTGCACGGGTGCTGCGCTGATGAGGCGCTGATGAGGCGCGCGCGCGCGCGCGCG 8441
QY 186 GCGCTTCTATAGGCGTTCTCCGCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 245
DB 8442 CGATGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8501
QY 246 CTTCGCCGCGTTCACGCGACGCGACTCTCCACCGAGCGCGCGCGCGCGCGCGCGCGCGCG 305
DB 8502 GCGGTCCACGGGACAGCGCGCGCGCTGCGGCGCTGCGGCGCTGCGGCGCGCGCGCGCG 8561
QY 306 TCCTCACCTCTCTCGACGACCTCTCTCGCGGGTTCAGGCGCGCGCGCGCGCGCGCGCG 365
DB 8562 GGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8621
QY 366 CAAGATCGGCGCATACGCTGGCCACCGAGTTTCGCGGAGCGCGCGCGCGCGCGCGCGCG 425
DB 8622 CGGGGACGGCTGACGACCGGGGACGCGCACCGTGGCGGCGCGCGCGCGCGCGCGCG 8681
QY 426 CGCCAAAGGACCGGGGAGCAGGAGCGTCTGCTGCGATTTCGCGGTTTCGCGGCGTTCG 485
DB 8682 CGGCGCGCGCTCGCGACCGCGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8741
QY 486 CGTGGCGCGCGGCGCGCGTGTGCGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 545
DB 8742 GCGCTCTCTGTAACCGCGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8801
QY 546 CGCGCGGCTCGCGCGCG 562
DB 8802 CGCGCGGACGGCGCTCG 8818

RESULT 10

US-09-105-537-5
; Sequence 5, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 36778
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-105-537-5

Query Match 4.2%; Score 57; DB 3; Length 36778;
Best Local Similarity 44.7%; Pred. No. 0.0016;
Matches 222; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

QY 66 CCGCGCGGAGCACCAGTTCG 125
DB 24009 CCGCGGACCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 24068
QY 126 CGACGGCTCCGCGCTCTTCTACAAAGCGCTCCAGGCGCGGAGCGCGCGCGCGCGCGCG 185
DB 24069 CGAGGCCCTGCACGGGTGCTGCGCTGATGAGGCGCTGCTGCGCGCGCGCGCGCGCG 24128
QY 186 GCGCTTCTATAGGCGTTCTCCGCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 245
DB 24129 CGATGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 24188

Qy	246	CTTCCCGGGTTCA	CGGGA	CGGACTCT	CCCA	CCGAGGCGAC	CCCGGGAGCCGA	305
Db	24189	CGGTTCCACGGG	CAGGCG	CCGCTT	GGGSCCT	CGCGGCT	CGCGCAGCGGAGCC	24248
Qy	306	TCCTCACTCGT	CTCTCG	ACACCTCT	CGCGGGTT	TCAGGCGCCT	TGGTGCAGACAT	365
Db	24249	GGGCGGTTCTG	TCTGCT	CGACT	TCGCGGGG	AGCCCG	CAGCGCGGGAGCCACGC	24308
Qy	366	CAGATCGGCGCA	TACGT	GGCCACCG	ATTC	CGCGGAGCC	TACATCGCCCAAGTGCT	425
Db	24309	CGGGACGGGCT	TACAG	CCGGGAC	CCAC	CCGTCGGCGG	CACCTCTGAGACGCGGCCT	24368
Qy	426	CGCCAAGGAC	CGGGGAC	CACGAG	CGCTT	CTGCTCGG	ATTCGCGCTCTCGCGCTCCGAGT	485
Db	24369	CGGACGGCCCT	CGAC	CGCCCT	CGGCT	CGGCGG	ACCGCAGCTCGCCCTCCGGGACGG	24428
Qy	486	CGTCGGCCCG	GAGGCG	CGCTGT	GGCGG	CAGAGCC	CCGAGGTGAAGGCATGGACAC	545
Db	24429	GGCGTCTCTG	TACCGG	CTTGG	CGGGG	CCCGCG	CGCCCGCGGCGGACGGCTTCG	24488
Qy	546	CGCGGGCGT	CCGCGG	562				
Db	24489	CGCGGCGG	ACGGCT	CG	24505			

```

RESULT 11
; US-09-105-537-1
; Sequence 1, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikomycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 15872
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
; US-09-105-537-1

```

Query Match	4.1%;	Score 55.2;	DB 3;	Length 15872;
Best Local Similarity	44.9%;	Pred. No. 0.0033;		
Matches	256;	Conservative	0;	Mismatches 308; Indels 6; Gaps 1;
QY	56	CGGACCTCCACCGCCGGAGCACCAAGTCGCCCGGCACCGCGCTCCCGCCAGCAAGCTGG	115	
Db	3849	CGGAGGCAACGCCACCCCGTCTCGCGTGATCGGGGACGGCGGTCAACACGAGACG	3908	
QY	116	GCCCGCTCATGACGGCTCGGGCTCTTTTACAAGCGCTTCAGAGCGCGGACGCTGGGG	175	
Db	3909	GCGCTTCCAACGGGTCAACGCCCCCAACGGCCGCTCCAGACGCGCTCATCGACAGG	3968	
QY	176	AGCAGAGTTCGCCTTCTATGAGCGTTCCTCCGCCACGCGCCGTCGCCGCCCGCATCC	235	
Db	3969	CCCTGGCCGACGCGCGGCTCAACCCCGCCGACGTCGACGCGCTGAGGGCGCACGGTACGG	4028	
QY	236	GAGACACCTTCTTCCCGGTTCCAGGCACGGGACTCTTCCCAACGAGGCGCAGCCCG	295	
Db	4029	GTACCCGGCTCGGCGACCCCATCGAGCGGAGCGGATCCTCGGCACCTACGGCGCGGACC	4088	
QY	296	GGGAGCGGCATCCTCACCTCGCTCTGACGACCTCTCTCGGGGTTTCAGGCGCCTCGG	355	
Db	4089	GGGGCGAGGGCGCTTCGCTCCAGCTCGGCTCGGTGAAGTCGAACATCGGCCACGCGCAGG	4148	
QY	356	TGCGAGACATCAAGATCGGCGGCATCAAGCTGGCCACCGAGTTTCGCGGAGCCCTACATCG	415	
Db	4149	CGGCCGCGGCGTGGGGGGCTCATAGATGGTCTTCGCGATGGCCACGCGTCTCTGC	4208	

Qy	416	CCAAAGTCCTCCCAAGAGACCGCGGGACCAAGAGCTTCTGCTCGGATTCGCGCTCTCCG	475
Db	4209	CCAGAGCGTCCACAGTGGACCGGCCACACCCGCG-----TCGACTGGAGAGCCGCGCG	4262
Qy	476	GCCTCCAGTTCGTGGCCCGCAGGGCCCGTGTGGCGACGGACGCCCGCGAGGTGAAGG	535
Db	4263	CGCTCGAGTCTCTACCGAGGAGCGGAGTGGCGGAGAGGGCGCGCCCGCGCGCGCGG	4322
Qy	536	CCATGGACACCGCGCGCGTCCGCGCGGTCTCCGGGCGCTACGTGTCTCATCGTTGCCGACG	595
Db	4323	CGATCTCTCTCTGGGATCAGCGGCACAAAGCCCATCTGTTGTCGAACAGGCCCCCGG	4382
Qy	596	AGGGATGGACTGTGCCTCGCCGCGCGG	625
Db	4383	AAGCCGGGAGCGCGGTCAACACACCG	4412

```

RESULT 12
US-09-091-609-1
; Sequence 1, Application US/09091609
; Patent No. 6600029
; GENERAL INFORMATION:
; APPLICANT: SHERMAN, DAVID H.
; APPLICANT: WILLIAMS, MARK D.
; APPLICANT: XUE, YONGQUAN
; TITLE OF INVENTION: METABOLIC ENGINEERING OF
; TITLE OF INVENTION: POLYHYDROXYALKANOATE MONOMER SYNTHASES
; FILE REFERENCE: 600.297US2
; CURRENT APPLICATION NUMBER: US/09/091.609
; CURRENT FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: PCT/US96/20119
; EARLIER FILING DATE: 1996-12-18
; EARLIER APPLICATION NUMBER: 60/008,847
; EARLIER FILING DATE: 1995-12-19
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 15872
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (20)...(13909)
US-09-091-609-1

```

Query Match	4.1%	Score 55.2	DB 4	Length 15872
Best Local Similarity	44.9%	Pred. No. 0.0033		
Matches 256	Conservative 0	Mismatches 308	Indels 6	Gaps 1
56	CCGACCTCCACCGCCGGAGACCAAGTCGCGGGCCACCGCGCTCCGCCACGACGAGCTGG	115		
3849	CCGAGCGCAACGGCCACCCCGTGCTCGCCGTGATCGGGGACGCGGCTCAACAGGACG	3908		
116	GCCCGCTCATGACGGCTCGGGCTTTCTCAAGCGGCTCCAGGCCGGACCGTGGGG	175		
3909	GCGCTCTCAACGGGCTACCGCCCCAACGGGCCCTCCACGACGGCGTCTATCCGACAGG	3968		
176	AGCAGGAGGTGCGCTTCTTAGAGGGGTTCTCGGCCACCGCGCGCTCCGGCCCGGCATCC	235		
3969	CCCTGGCGCAGCGCGGGCTACCCCGGCCGAGCTCGACGCCGTGAGGCGGCAGCTACGG	4028		
236	GAGACACTTCTTCCCGCGTTCACGGACGCGACTCTCCGCCACCGAGGCGAGCCCG	295		
4029	GTACCGCGTTCGCGACCCCATCGAGGGCGAGGGGATCTCTGGCACCTTACGGCGGGACC	4088		
296	GGAGCGCGCATCTTCACTTCGTCTTCGACGACTCTCTCGGGGGTTTCAGSGCGCCTGCG	355		
4089	GGGGCGAGGGCGCTCGGTTCCAGTCCGCTCGCTGAAGTGGAAATCGGCCACCGCGCAGG	4148		
356	TGSCAGACATCAAGATCGCGGCCATACGTGGCCACCGAGTTTCGCGGAGCGCTTACATCG	415		
4149	CCGCGCGGGCGTGGCGGGGCTCATCAGATGTGTCCTCGGATCGGCCACGGGGTCTCTGC	4208		

QY 416 CCAAGTGCCTCGCCAGGACCGCGGACCAAGAGCGTTCTGCTCGGATTCCGGTCTCCG 475
Db |||||
QY 4209 CCAGGACGCTCCACGTGAGCCGCGCCACCAACCCGCG-----TCGACTGGAGGCGCGCG 4262
Db |||||
QY 476 GCCTCCGAGTCTCGGCGCCCGAGGCGCGCTGTGCGGAGCGAGCGCCCGGAGTGAAG 535
Db |||||
QY 4263 GCGTCGAGCTCTCACCGAGGAGCGGAGTGGCCGAGAGCGGCGCCCGCGCGCGCG 4322
Db |||||
QY 536 CCATGGACACCGCGCGCTCGCGCGCTGTGCGGCGCTACGTGTCTATCCGTTTGGCGACG 595
Db |||||
QY 4323 CGATCTCTCTTCGCGCATCAGCGGACCAAGCGCCACATCGTGGTGGTGAACAGGCGCCCG 4382
Db |||||
QY 596 AGGGATGAGTGTGCGCTCGCGCGCGCGG 625
Db |||||
QY 4383 AGCGGGGAGGCGCGGTCAACACACCG 4412
Db |||||

RESULT 13

US-09-091-609-3
; Sequence 3, Application US/09091609
; Patent No. 6600029
; GENERAL INFORMATION:
; APPLICANT: SHERMAN, DAVID H.
; APPLICANT: WILLIAMS, MARK D.
; APPLICANT: XUE, YONGQUAN
; TITLE OF INVENTION: METABOLIC ENGINEERING OF
; POLYHYDROXYALKANOATE MONOMER SYNTHASES
; FILE REFERENCE: 600.297052
; CURRENT APPLICATION NUMBER: US/09/091,609
; EARLIER APPLICATION NUMBER: PCT/US96/20119
; EARLIER FILING DATE: 1996-12-18
; EARLIER APPLICATION NUMBER: 60/008,847
; EARLIER FILING DATE: 1995-12-19
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 15872
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (14148)....(15824)
US-09-091-609-3

Query Match 4.1%; Score 55.2; DB 4; Length 15872;
Best Local Similarity 44.9%; Pred. No. 0.0033;
Matches 256; Conservative 0; Mismatches 308; Indels 6; Gaps 1;
QY 56 CCGACTCCACCGCGCGGAGACCAAGTCCGCGGCCACCGCGCTTCGCGCAAGCTGG 115
Db |||||
QY 3849 CCGAGCGCAACGCCACCCCGCTCGCGTGTATCCGGGCGAGCGGGTCAACGAGGAG 3908
Db |||||
QY 116 GCGCGTATACAGCGCTCGCGCTCTCTAAGCGCTCCAGGCGCGGACCGTGGG 175
Db |||||
QY 3909 GCGCTCCAGCGGTACCGCGCCCAAGCGCGCTCCAGAGCGCGTCACTCCGACAG 3968
Db |||||
QY 176 AGCAGAGTCCCTTCTATGAGGCGTCTCTCGCCACCGCGCGCTCCGCGCCGCATCC 235
Db |||||
QY 3969 CCTGCGGACCGCGGCTACCGCGCGAGTGAACCGCTTCGAGGCGACGATACG 4028
Db |||||
QY 236 GAGACACTTCTTCCCGGTTCCAGGCAACGCGACTCTCTCCCGCCAGGCGCGAGCCG 295
Db |||||
QY 4029 GTACCGGCTCGCGACCCCATCGAGGCGAGGCGATCTCTGCGACCTACGCGCGGACC 4088
Db |||||
QY 296 GGGAGCGCATCTCACTCGTCTCGAGCACTCTCTCGCGGGTTTACGGCGCCCTGCG 355
Db |||||
QY 4089 GGGGAGGCGCTCGCTTCAGCTCGGCTCGGCTGAAATCGAATCGGCCACGCGGAG 4148
Db |||||
QY 356 TCGCAGACATCAAGATCGCGCGCATCAGTGGCCACCGAGTTTCGCGGAGCCCTACATCG 415
Db |||||
QY 4149 CCGCGCGGCGTGGCGGGCTCATCAGATGTTCTCGCGATGCGCCACGCGCTCTCG 4208
Db |||||

QY 416 CCAAGTGCCTCGCCAGGACCGCGGACCAAGAGCGTTCTGCTCGGATTCCGGTCTCCG 475
Db |||||
QY 4209 CCAGGACGCTCCACGTGAGCCGCGCCACCAACCCGCG-----TCGACTGGAGGCGCGCG 4262
Db |||||
QY 476 GCCTCCGAGTCTCGGCGCCCGAGGCGCGCTGTGCGGAGCGAGCGCCCGGAGTGAAG 535
Db |||||
QY 4263 GCGTCGAGTCTCTCACCGAGGAGCGGAGTGGCCGAGAGCGGCGCCCGCGCGCGCG 4322
Db |||||
QY 536 CCATGGACACCGCGCGCTCGCGCGCTGTGCGGCGCTACGTGTCTATCCGTTTGGCGACG 595
Db |||||
QY 4323 CGATCTCTCTTCGCGCATCAGCGGACCAAGCGCCACATCGTGGTGGTGAACAGGCGCCCG 4382
Db |||||
QY 596 AGGGATGAGTGTGCGCTCGCGCGCGCGG 625
Db |||||
QY 4383 AGCGGGGAGGCGCGGTCAACACACCG 4412
Db |||||

RESULT 14

US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 4.1%; Score 55.2; DB 3; Length 4403765;
Best Local Similarity 43.4%; Pred. No. 0.014; Mismatches 328; Indels 0; Gaps 0;
Matches 252; Conservative 0; Mismatches 328; Indels 0; Gaps 0;
QY 46 CATACATGTCCGACTCCACCGCGCGGAGCACCAGTCCGCGGCCACCGCGCTCCGCG 105
Db |||||
QY 150482 CAGCCACCGCGCGCACCGCGGCGCCATGCTGCTGACCGGCGAGGCGGCGCC 150423
Db |||||
QY 106 AGCAGCTGGCGCGCTCATCGAGCTCGCGCTCTTCTACAGCGGTCCAGGCGCGC 165
Db |||||
QY 150422 GTACACACGCTTTCGCGCGCTGTCGCGCGCGCGGCGAGTACCCCAAGTCCGCG 150363
Db |||||
QY 166 GACCTGGGAGCAGCAGTCTGCTTCTATGAGGCGTTCGCGCCACGCGCGCTCCG 225
Db |||||
QY 150362 GCGCGCGCTTGGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTGTCC 150303
Db |||||
QY 226 GCGCGCATCCGAGACACTTCTTCCCGCGGTTCACAGCGCAGCGACTCTCTCCCAACCGAG 285
Db |||||
QY 150302 GCGGTGTCGCGCGCGCGCGCTGCGCGTAGAGCATCCGCGCACCGCGCACCGCGC 150243
Db |||||
QY 286 GCGCAGCCCGGAGCGCGCTCTTCTCTCTGAGCAGCTCTCTCGCGGGTTTCA 345
Db |||||
QY 150242 GAGGCCACTCTCTCCAGCTGACGATGTTGGCATCGCGCGCTTCCGCGCATTCGCCAC 150183
Db |||||
QY 346 GCGCGCTGCTGCGCAGACATCAAGATCGCGCGCGCATCAGTGGCGCACCGAGTTTCGCGGAG 405
Db |||||
QY 150182 GCG 150123
Db |||||
QY 406 CCCTACATCGCCAAAGTGCCTTCGCGCAAGAGCGCGGACCAAGAGCGGCTTCTGCTCGGATTC 465
Db |||||

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2004, 03:33:25 ; Search time 513 Seconds
(without alignments)
9753.850 Million cell updates/sec

Title: US-10-042-894A-7
Perfect score: 1344
Sequence: 1 gcacgaggtcagtcgctcac.....ataaaaaaaaaaaaaa 1344

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2458946 seqs, 1861504846 residues

Total number of hits satisfying chosen parameters: 4517892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1344	100.0	1344	14	US-10-042-894A-7
2	865.4	64.4	923	14	US-10-042-894A-5
3	844.8	62.9	1169	14	US-10-042-894A-1
4	837.4	62.3	923	14	US-10-042-894A-3
5	740.4	55.1	3416	14	US-10-042-894A-20
6	686	51.0	899	14	US-10-042-894A-15
7	470.4	35.0	643	14	US-10-042-894A-17
8	298.4	22.2	519	14	US-10-042-894A-18
9	191.6	14.3	353	14	US-10-042-894A-19
10	184.6	13.7	1195	14	US-10-042-894A-11
11	179	13.3	1105	14	US-10-042-894A-9
12	149.8	11.1	1020	14	US-10-042-894A-13
13	149.8	11.1	1020	14	US-10-042-894A-13
14	104.6	7.8	464	9	US-09-770-444-340
15	71.4	5.3	135638	14	US-10-314-657-1

68.6 5.1 1383 12 US-10-417-700A-10
17 68.6 5.1 88421 9 US-09-976-059-1
18 68.2 5.1 2541 14 US-10-156-761-2855
19 68.2 5.1 9025608 14 US-10-156-761-1
20 66.8 5.0 1347 12 US-10-424-599-118038
21 66.2 4.9 3471 14 US-10-156-761-3726
22 65.4 4.9 484 10 US-09-770-961-829
23 64.8 4.8 825 14 US-10-156-761-4706
24 64.8 4.8 9025608 14 US-10-156-761-1
25 64.2 4.8 1296 12 US-10-425-114-33515
26 63.8 4.7 5452 14 US-10-017-161-1481
27 63.8 4.7 5452 15 US-10-292-798-1189
28 63.6 4.7 987 14 US-10-156-761-3940
29 63.4 4.7 1581 14 US-10-156-761-3431
30 61.8 4.6 7788 14 US-10-329-079-8
31 61.8 4.6 11058 14 US-10-156-761-3629
32 61.8 4.6 37360 14 US-10-329-079-6
33 61.4 4.6 18435 14 US-10-156-761-412
34 61.4 4.6 100000 14 US-10-260-238-5598
35 61.2 4.6 2475 15 US-10-425-114-16619
36 61.2 4.5 2662 12 US-10-425-114-4294
37 61 4.5 1271 12 US-10-369-493-42318
38 61 4.5 1385 12 US-10-378-083-1
39 61 4.5 2466 15 US-10-145-413-104
40 61 4.5 64492 12 US-10-425-114-16568
41 60.2 4.5 1266 14 US-10-425-114-27827
42 60.2 4.5 2291 12 US-10-425-114-28844
43 60 4.5 1485 12 US-10-156-761-6190
44 60 4.5 1500 12 US-10-156-761-6190
45 59.8 4.4 1440 14

ALIGNMENTS

RESULT 1

US-10-042-894A-7
; Sequence 7, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase
; TITLE OF INVENTION: Genes and Uses Thereof
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1344
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52)...(921)
US-10-042-894A-7

Query Match 100.0%; Score 1344; DB 14; Length 1344;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCACGAGGTGAGTCCGTCACCCCTCGGCCCATAGTCCCTTCCCATACCATGTCGCAC 60
Db 1 GCACGAGGTGAGTCCGTCACCCCTCGGCCCATAGTCCCTTCCCATACCATGTCGCAC 60
QY 61 CTCACCCGCGGAGGACCAAGTCGCGGCCACCGCGCTCCGCGAGCTGGGCGG 120

Db	61	CTCACC	CGCGGAGCACAAGTCGCGGCGCACCGCGCTCCGCCACGACAAAGCTGGCGCGG	120
Qy	121	CTCATCGAGG	CTCGGCGCTCTTCTACAAAGCCGCTCCAGCGCGGCGACGCTGGGAGCAC	180
Db	121	CTCATCGAGG	CTCGGCGCTCTTCTACAAAGCCGCTCCAGCGCGGCGACGCTGGGAGCAC	180
Qy	181	GAGGTGCGCTT	CTATGAGCGGTTCTCGGCCACGCGCGCGTCCCGCGCGCATCCGAGAC	240
Db	181	GAGGTGCGCTT	CTATGAGCGGTTCTCGGCCACGCGCGCGTCCCGCGCGCATCCGAGAC	240
Qy	241	ACCTTCTT	CCCCCGGTTCACGCGCACGGGACTCCCTCCACAGCGGCGCAGCCCGGGGAG	300
Db	241	ACCTTCTT	CCCCCGGTTCACGCGCACGGGACTCCCTCCACAGCGGCGCAGCCCGGGGAG	300
Qy	301	CGCATCTCA	CTCTCTCGACACTCTCTCGCGGGGTTTCAGCGCGCTTCGCGCTCGCA	360
Db	301	CGCATCTCA	CTCTCTCGACACTCTCTCGCGGGGTTTCAGCGCGCTTCGCGCTCGCA	360
Qy	361	GACATCAAGAT	TCGGGCCATCAGTGGGCAACGAGTTCGCGGAGCCCTACATGCCCAAG	420
Db	361	GACATCAAGAT	TCGGGCCATCAGTGGGCAACGAGTTCGCGGAGCCCTACATGCCCAAG	420
Qy	421	TGCCTCGC	AAAGGACCGGGGACACAGAGCGTTCGTCGGATTCGCGCTTCGCGCGTC	480
Db	421	TGCCTCGC	AAAGGACCGGGGACACAGAGCGTTCGTCGGATTCGCGCTTCGCGCGTC	480
Qy	481	CGAGTCGT	CGCCCGGAGGCGCGTGTGGGCGACGGAGCGCCCGAGGTGAAGCCCATG	540
Db	481	CGAGTCGT	CGCCCGGAGGCGCGTGTGGGCGACGGAGCGCCCGAGGTGAAGCCCATG	540
Qy	541	GACACCGCG	CGTTCGCGCGTGTTCGCGCGTACGTTTCGCGCGAGGGG	600
Db	541	GACACCGCG	CGTTCGCGCGTGTTCGCGCGTACGTTTCGCGCGAGGGG	600
Qy	601	ATGACTGT	CGCTCGCGCGCGGTGTACCGAGAAAGTGGAGTTCTGTCAAGCTG	660
Db	601	ATGACTGT	CGCTCGCGCGCGGTGTACCGAGAAAGTGGAGTTCTGTCAAGCTG	660
Qy	661	CGGAGCTC	AAGGCGTGTTCGAGGAGCAGACTCTGTTCACCTTCTACTCGGCGTCAAT	720
Db	661	CGGAGCTC	AAGGCGTGTTCGAGGAGCAGACTCTGTTCACCTTCTACTCGGCGTCAAT	720
Qy	721	CTTCTGGG	CTATGATGTCGTGAGTCGACGCGGAGATGGGGTGGGTGACGCTG	780
Db	721	CTTCTGGG	CTATGATGTCGTGAGTCGACGCGGAGATGGGGTGGGTGACGCTG	780
Qy	781	AAGCTGTG	GACTTTCGCTATGTCGCGAGGGTGTGGGTGATGGGTGACGCTG	840
Db	781	AAGCTGTG	GACTTTCGCTATGTCGCGAGGGTGTGGGTGATGGGTGACGCTG	840
Qy	841	GGCGGCTC	TCGCTGATCAAGTTCGATTCGACATTTCTGACATTTTCGCGAGACTTCCTATACG	900
Db	841	GGCGGCTC	TCGCTGATCAAGTTCGATTCGACATTTCTGACATTTTCGCGAGACTTCCTATACG	900
Qy	901	CAGGCTTT	GGTCTCTTAAGAGGATCTTGCAATTCGATTTGATAACAAAGCCCTA	960
Db	901	CAGGCTTT	GGTCTCTTAAGAGGATCTTGCAATTCGATTTGATAACAAAGCCCTA	960
Qy	961	CAAGTTT	GTCTGAAAAGAGCGCTCGAGTTGTGGGTGTGGAGATCTGAGACG	1020
Db	961	CAAGTTT	GTCTGAAAAGAGCGCTCGAGTTGTGGGTGTGGAGATCTGAGACG	1020
Qy	1021	GTGTCGG	CGCCTTGTGCTTTGCCCTTTGCCCTGCTGCAAAACATACGCAACCTGCT	1080
Db	1021	GTGTCGG	CGCCTTGTGCTTTGCCCTTTGCCCTGCTGCAAAACATACGCAACCTGCT	1080
Qy	1081	CCTTTTT	TCGAAACCCCTTACTTCGGAAGAAACTTTTTTTTCCCACTTTGGGGGTTCGA	1140
Db	1081	CCTTTTT	TCGAAACCCCTTACTTCGGAAGAAACTTTTTTTTCCCACTTTGGGGGTTCGA	1140
Qy	1141	TTACGTT	TGGATCTGTTGTGGCACTCGGTTCAGAGTTGTAAAGATCGAGGGAGCGGTG	1200
Db	1141	TTACGTT	TGGATCTGTTGTGGCACTCGGTTCAGAGTTGTAAAGATCGAGGGAGCGGTG	1200

RESULT 2

```

US-10-042-894A-5
; Sequence 5, Application US/10042894A
; Publication No. US2003009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongsyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cancon, Rebecca E.
; TITLE OF INVENTION: No. US2003009011
; TITLE OF INVENTION: Genes and Uses
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042-894A-5
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ IDS NOS: 37
; SOFTWARE: FastSeq for Windows Version 1.0
; SEQ ID NO 5
; LENGTH: 923
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (53)...(922)
US-10-042-894A-5

```

Query Match 64.4%; Score 865.4; DB 14; Length 923;
Best Local Similarity 98.8%; Pred. No. 1.1e-222;
Matches 870; Conservative 0; Mismatches 11; Indels 0;

42	QY	TCCCCATACATGTCGGA	CTCCACCGCGCGAGAC	CCAAAGTCGCCGGCCACCGCGCTC	101
43	Db	TCCCCATACATGCCGAC	CTCCACCGCGGAGAC	CCAAAGTCGCGGTACCGCGCTC	102
102	QY	CGCAGCAGCTGGGCGCGT	CATCGACGCTCGGCTCTT	CACAAGCGGCTCGAGC	161
103	Db	CGCAGCAAGCGGGCGCGT	CATCGACGCTCGGCTCTT	CACAAGCGGCTCGAGC	162
162	QY	CGGCGACCGTGGGAG	CACGAGGTCGGCTTCTAT	GAGCGGTTCTCGCCGCCACCGCGCGT	221
163	Db	CGGCGACCGTGGGAG	CACGAGGTCGGTCTTAT	GAGCGGTTCTCGCCGCCACCGCGCGT	222
222	QY	CCCGGCCGCATCCGAG	CACCTTCTTCCCGCGTTC	CACGGCAGCGACTCTCTCCCGAC	281
223	Db	CCCGGCCGCATCCGAG	CACCTTCTTCCCGCGTTC	CACGGCAGCGACTCTCTCCCGAC	282
282	QY	CGAGGCGAGCCCGGG	AGCGGATCCTCACTCGTCT	CGACGACTCTCGCGGGGTT	341
283	Db	CGAGGCGAGCCCGGG	AGCGGATCCTCGACCTCT	CGACGACTCTCTCGCGGGATT	342
342	QY	TGAGGCGCCCTGGT	CGCAGACATCAAGATCG	GCGCATCACTGGGCCACCGAGTTCCGC	401
343	Db	TGAGGCGCCCTGGT	CGCAGACATCAAGATCG	GCGCATCACTGGGCCACCGAGTTCCGC	402
402	QY	GGAGCCCTACATCGC	CAAGTGGCTCCGCCAAGG	ACCGGGACACCGAGCGTCTTGCTCGG	461
403	Db	GGAGCCCTACATCGC	CAAGTGGCTCCCTTG	ACCGGGACACCGAGCGTCTTGCTCGG	462

;; PRIOR APPLICATION NUMBER: US 60/261,465
;; PRIOR FILING DATE: 2001-01-12
;; NUMBER OF SEQ ID NOS: 37
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 3
;; LENGTH: 923
;; TYPE: DNA
;; ORGANISM: Zea mays
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (53)...(736)
US-10-042-894A-3

Query Match 62.3%; Score 837.4; DB 14; Length 923;
Best Local Similarity 97.5%; Pred. No. 3.8e-215;
Matches 859; Conservative 0; Mismatches 21; Indels 1; Gaps 1;
QY 42 TCCCATACCATGTCGACCTTCCACCCGCGGAGCACCAGTCCGCGGCCACCGCGCTC 101
DB 43 TCCCATACCATGTCGACCTTCCACCCGCGGAGCACCAGTCCGCGGCCACCGCGCTC 102
QY 102 CGCAGCAAGTGGGCGGCTCATCGAGGCTCCGGCTCTTCTACAGCGGCTCCAGGC 161
DB 103 CGCAGCAAGTGGGCGGCTCATCGAGGCTCCGGCTCTTCTACAGCGGCTCCAGGC 162
QY 162 CGCGACCGTGGGAGCAGGAGTCCGCTTCTATGAGCGTTCCTCCGCCACGCGCGCT 221
DB 163 CGCGACCGTGGGAGCAGGAGTCCGCTTCTATGAGCGTTCCTCCGCCACGCGCGCT 222
QY 222 CCGCGCGCATCCGAGACACTTCTTCCCGCGTTCGAGGACGCGACTCTCCCGAC 281
DB 223 CCGCGCGCATCCGAGACACTTCTTCCCGCGTTCGAGGACGCGACTCTCCCGAC 282
QY 282 CGAGCGGAGCGCGGAGCGGCTCTCTCACTCTCTCGACGACTCTCCGGGAGTT 341
DB 283 CGAGCGGAGCGCGGAGCGGCTCTCTCACTCTCTCGACGACTCTCCGGGAGTT 342
QY 342 TCAGCGCGCTCGCTCGCAGACATCAAGATCGGCGGCATCACTGCGGCACCGAGTTCGCC 401
DB 343 TGAGCGCGCTCGCTCGCAGACATCAAGATCGGCGGCATCACTGCGGCACCGAGTTCGCC 402
QY 402 GGAGCGCTTACATCGCCAGTCCCTCGCCAGAGCCGCGGACCGAGCGCTTCGTCGG 461
DB 403 GGAGCGCTTACATCGCCAGTCCCTCGCCAGAGCCGCGGACCGAGCGCTTCGTCGG 462
QY 462 ATTCCGGCTCTCGCGCTCGAGTCTGTCGGCCCGAGGCGCGCTGTGGCGAGCGAGCG 521
DB 463 ATTCCGGCTCTCGCGCTCGAGTCTGTCGCCGCGCGCGCTGTGGCGAGCGAGCG 522
QY 522 CCGGAGGTGAAGGCATGGACACCGCGCGCTCGCGCGCTGCTCCGCGCTACGTGTC 581
DB 523 CCGGAGGTGAAGGCATGGACACCGCTCGGCTCGCGCGCTGCTCCGCGCTACGTGTC 582
QY 582 ATCCG-TTGCAGAGGAGTGAAGTGTGCTGCTCGCGCGCGGCTGACGAGGAGAAAG 640
DB 583 ATCCG-TTGCAGAGGAGTGAAGTGTGCTGCTCGCGCGCGGCTGACGAGGAGAAAG 642
QY 641 GTGGAGTCTTGTCAAGCTCGCGAGCTCAAGGCTGCTCGAGGAGCAGACTCTGTTC 700
DB 643 GTGGAGTCTTGTCAAGCTCGCGAGCTCAAGGCTGCTCGAGGAGCAGACTCTGTTC 702
QY 701 ACTTCTACTCGCGCTCGATTTCTTGGGCTATGATGCTGCTGCACTCGCAGCAGCGGAG 760
DB 703 ACTTCTACTCGCGCTCGATTTCTTGGGCTATGATGCTGCTGCACTCGCAGCAGCGGAG 762
QY 761 ATGGGGTGGGTGACGAGTGAAGTGTGAGATTTTGGCCATGTCGCCAGGAGGTGATGGG 820
DB 763 GTGGGGTGGGTGAACAGTGAAGTGTGAGATTTTGGCCATGTCGCCAGGAGGTGATGGG 822
QY 821 TGATTTACCAACAATCTTGGCGGCTCTGCTGCTGATCAAGTTCGTTCTGACATTG 880
DB 823 TGATTTACCAACAATCTTGGCGGCTCTGCTGCTGATCAAGTTCGTTCTGACATTG 882

QY 881 TTCCGAGAGCTCCTCATAGCGAGCCCTTGGGTCTTCTTAA 921
DB 883 TTCCGAGAGCTCCTCATAGCGAGCCCTTGGGTCTTCTTAA 923

RESULT 5

US-10-042-894A-20
;; Sequence 20, Application US/10042894A
;; Publication No. US20030009011A1
;; GENERAL INFORMATION:
;; APPLICANT: Shi, Jinrui
;; APPLICANT: Beach, Larry
;; APPLICANT: Wang, Hongyu
;; APPLICANT: Rafalski, Antoni J.
;; APPLICANT: Cahoon, Rebecca E.
;; TITLE OF INVENTION: No. US20030009011A1 Inositol Polyphosphate Kinase
;; TITLE OF INVENTION: Genes and Uses Thereof
;; FILE REFERENCE: 1286
;; CURRENT APPLICATION NUMBER: US/10/042,894A
;; CURRENT FILING DATE: 2002-01-09
;; PRIOR APPLICATION NUMBER: US 60/261,465
;; PRIOR FILING DATE: 2001-01-12
;; NUMBER OF SEQ ID NOS: 37
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 20
;; LENGTH: 3416
;; TYPE: DNA
;; ORGANISM: Zea mays
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (72)...(407)
US-10-042-894A-20

Query Match 55.1%; Score 740.4; DB 14; Length 3416;

Best Local Similarity 90.2%; Pred. No. 6.4e-189;
Matches 862; Conservative 0; Mismatches 31; Indels 63; Gaps 4;

QY 2 CACGAGTCACTGCTGTCACCCCTCGCGCCCATAGTCCCTTCCCATACCATGTCGCGACC 61
DB 28 CACCATCGCCACCGTTCACCCCTTGTCCCATAG-----TCCCATACCATGTCGCGACC 81
QY 62 TCACCCGCGGAGCACCAGTCCGCGCCACCGCGCTCGGCGAGCAAGCTGGCGCGC 121
DB 82 TCACCCGCGGAGCACCAGTCCGCGCTCACCGCGCTCGGCGAGCAAGCTGGCGCAC 141
QY 122 TCATCGAGCGCTCGGCGCTCTTACAGCGCTTCCAGCGCGCGACCGTGGGAGCAG 181
DB 142 TCATCGAGCGCTCGGCGCTCTTACAGCGCTTCCAGCGCGCGACCGTGGGAGCAG 201
QY 182 AGTTCGCTTCTATGAGCGCTTCTCGCCACCGCGCGCTCGCGCGCGCATCCGAGACA 241
DB 202 AGTTCGCTTCTATGAGCGCTTCTCGCCACCGCGCGCTCGCGCGCGCATCCGAGACA 261
QY 242 CTTCTTTCGCGGTTCCAGCGACCGAGCTCTTCCCGACCGAGCGCGCGCGGAGC 301
DB 262 CTTCTTTCGCGGTTCCAGCGACCGAGCTCTTCCCGACCGAGCGCGCGCGGAGC 321
QY 302 CGCATCTCTACCTCGTCTCTGAGCGACCTCTCGCGGGTTTCAGGCGCGCTTGTGCGAG 361
DB 322 CGCATCTCTACCTCGTCTCTGAGCGACCTCTCGCGGGTTTCAGGCGCGCTTGTGCGAG 381
QY 362 ACATCAAGATCGGCGCATCATCGTGGCCACCGAGTTCGCGGAGCCCTTACATCCCAAGT 421
DB 382 ACATCAAGATCGGTCGCATCACT----- 405
QY 422 GCCTCGCAAGGACCGCGGAGCAGCGCTTCTGCTCGGATTCGCGCTCTCCGCGCTCC 481
DB 406 -----GACCAAGAGCGTTCGCTCGGATTCGCGCTCTCCGCGCTCC 446
QY 482 GAGTCTCGGCGCGAGCGCGCTGTGGCGAGCGGAGCGCGCGGAGGTGAAGCCATGG 541
DB 447 GAGTCTCGGCGCGAGCGCGCTGTGGCGAGCGGAGCGCGCGGAGGTGAAGCTATGG 506


```
QY 542 ACACCGCGCGTCCGCGCGTGTCTCCGGCGCTACGTGTATCGG-TTGGCGAGAGGGG 600
Db 507 ACATTGTCCGGGTCCGCGCGTGTCTCCGGCGCTACGTGTATCGGCTTGGCGAGAGGG 566
QY 601 ATGGACTGTGCGTCCGCGCGGTGTACGAGAGAAAGTGGAGTCTTGTCAAGCTG 660
Db 567 ATGGACTGTGCGTCCGCGCGGTGTACGAGAGAAAGTGGAGTCTTGTCAAGCTG 626
QY 661 CGGAGCTCAAGGCTGGTTCGAGGAGAGACTGTCTTGTCACTTCTACTCGGCGTCAATT 720
Db 627 CGGAGCTCAAGGCTGGTTCGAGGAGAGACTGTCTTGTCACTTCTACTCGGCGTCAATT 686
QY 721 CTTCTGGGCTATGATGCTGCTGCAATCGCAGAGCGAGATGGGGGTGGGGTAAAGTG 780
Db 687 CTTCTGGGCTATGATGCTGCTGCAATCGCAGAGCGAGAGTGGGGTAAAGTG 746
QY 781 AAGTGTGTGACTTTGGCCATGTGGCGAGAGTGTGGGGTGAATGACCAACTTCTCTG 840
Db 747 AAGTGTGTGACTTTGGCCATGTGGCGAGAGTGTGGGGTGAATGACCAACTTCTCTG 806
QY 841 GGCGGGCTCTGCTCGCTGATCAAGTGTCTTGTCACTTGTTCGAGAGTCTCTCAACG 900
Db 807 GGCGGGCTCTGCTCGCTGATCAAGTGTCTTGTCACTTGTTCGAGAGTCTCTCAACG 866
QY 901 CAGCCTTTGGGTCTCTTAAAGAGAGGATCCTGGCA-TTTCGATTTCATAACAAAG 955
Db 867 CAGCCTTTGGGTCTCTTAAAGAGAGGATCCTGGCA-TTTCGATTTCATAACAAAG 922

RESULT 6
US-10-042-894A-15
; Sequence 15, Application US/10042894A
; Publication No. US2003009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US2003009011A1 Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; PRIOR FILING DATE: 2002-01-09
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 899
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (89) ... (424)
US-10-042-894A-15

Query Match 51.0%; Score 686; DB 14; Length 899;
Best Local Similarity 88.7%; Pred. No. 2e-174;
Matches 802; Conservative 0; Mismatches 40; Indels 62; Gaps 3;

QY 2 CACGAGGTCACTGCTGCTACCCCTCGCGCCATAGTCCCTTCCCATACCAATGTCGACG 61
Db 45 CACATCGCCACCGTGTACCCCTTGTCTCCCATAG-----TCCCATACCAATGTCGACG 98
QY 62 TCCACCCCGCGAGACCAAGTTCGCGGCCACCGCGCTCCGCCAGCAAGTGGGCGCGC 121
Db 99 TCCACCCCGCGAGACCAAGTTCGCGGCCACCGCGCTCCGCCAGCAAGTGGGCGCGC 158
QY 122 TCATCGAGGCTCGGCGCTCTTCTACAGCGCTTCCAGCGCGGAGCGTGGGAGACG 181
Db 159 TCATCGAGGCTCGGCGCTCTTCTACAGCGCTTCCAGCGCGGAGCGTGGGAGACG 218
```

```
QY 182 AGCTCGCTTCTATGAGGCGTCTTCGCGCCACGCGCGCTCCGCGCCGATCCGAGACA 241
Db 219 AGTTCGCTTCTATGAGGCGTCTTCGCGCCACGCGCGCTCCGCGCCGATCCGAGACA 278
QY 242 CTTCTTCCCGCGTTCACCGCAAGCTCTTCCCAACCGAGGCGCAAGCGGGGAGC 301
Db 279 CTTCTTCCCGCGTTCACCGCAAGCTCTTCCCAACCGAGGCGCAAGCGGGGAGC 338
QY 302 CGCATCTCACTCTGCTTCGAGAGACTCTCTCGGGGGTTCAGGCGCGCTCGCTCCGAG 361
Db 339 CGCATCTCACTCTGCTTCGAGAGACTCTCTCGGGGGTTCAGGCGCGCTCGCTCCGAG 398
QY 362 ACATCAAGATCGCGCGCATCACGTGGCCACCGAGTTCGCGGAGCCCTACATCCCAAGT 421
Db 399 ACATCAAGATCGGTGCTCATCATCT----- 422
QY 422 GCCTCGCAAGGACCGCGGACCAAGAGCTTCTCTCGGATTCGCGCTCTCGGCGCTCC 481
Db 423 -----GACCATGAGCGATCTGCTCGGATTCACGCTCTCGGCGCTCC 463
QY 482 GAGTCTCGCGCGCGGCGCGTGTGGCGGAGCGCGGAGGCGCGGAGGCGCGGAGGCG 541
Db 464 GAGTCTCGCGCGCGGCGCGTGTGGCGGAGCGCGGAGGCGCGGAGGCGCGGAGGCG 523
QY 542 ACACCGCGCGCTCCGCGCGTGTCTCGGCGCTACGTGTCTCATCCG-TTGCCTGAGAGGG 600
Db 524 ACATTGTGCGCGTCCGCGCGTGTCTCGGCGCTGCTCATCGCTTTCGCGCGAGGG 583
QY 601 ATGAGCTGTGCTGCGCGCGGTGTACGAGGAGAAAGGTGGAGTCTTGTCAAGCTG 660
Db 584 ATGAGCTGTGCTGCGCGCGGTGTACGAGGAGAAAGGTGGAGTCTTGTCAAGCTG 643
QY 661 CGGAGCTCAAGGCGGTGTTCGAGGAGCAGACTCTTCCACTTCTACTCGGCGTCAATT 720
Db 644 CGGAGCTCAAGGCGGTGTTCGAGGAGCAGACTCTTCCACTTCTACTCGGCGTCAATT 703
QY 721 CTTCTGGGCTATGATGCTGCTGCAAGTTCGAGGCGGAGATGGGGTGGGGTGAACAGTG 780
Db 704 CTTCTGGGCTATGATGCTGCTGCAAGTTCGAGGCGGAGTGGGGTGGGGTGAACAGTG 763
QY 781 AAGCTGTGAGCTTGGCCATGTGGCGAGGCGGAGTGGGGTGAACCAACTTCTCTG 840
Db 764 AAGCTGTGAGCTTGGCCATGTGGCGAGGCGGAGTGGGGTGAACCAACTTCTCTG 823
QY 841 GGCGGGCTCTGCTGCTGATCAAGTTCGAGTTCGAGTTCGAGGAGCTTCCTATACG 900
Db 824 GGCGGGCTCTGCTGCTGATCAAGTTCGAGTTCGAGTTCGAGGAGCTTCCTATACG 883
QY 901 CAGC 904
Db 884 CAGC 887

RESULT 7
US-10-042-894A-17
; Sequence 17, Application US/10042894A
; Publication No. US2003009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US2003009011A1 Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; PRIOR FILING DATE: 2002-01-09
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 17
```

```

; LENGTH: 643
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(643)
; OTHER INFORMATION: n = A, T, C or G
US-10-042-894A-17

Query Match      35.0%; Score 470.4; DB 14; Length 643;
Best Local Similarity 95.9%; Pred. No. 1.8e-116;
Matches 509; Conservative 0; Mismatches 18; Indels 4; Gaps 3;

QY 42 TCCCCATACCATGTCGACCTCACCGCGCGGAGACCAAGTCCGCGGCCACCGCGCCCTC 101
Db 107 TCCCCATACCATGTCGACCTCACCGCGCGGAGACCAAGTCCGCGGCCACCGCGCCCTC 166
QY 102 CGCCAGCAAGCTGGCGCGCTCATGACAGGCTCCGGCTCTTCTACAAGCCGCTCCAGGC 161
Db 167 CGCCAGCAAGCTGGCGCGCTCATGACAGGCTCCGGCTCTTCTACAAGCCGCTCCAGGC 226
QY 162 CGCGACCGTGGGAGCACGAGTTCGCTTCTATGAGGGTCTTCCGCCACCGCCCGT 221
Db 227 CGCGACCGTGGGAGCACGAGTTCGCTTCTATGAGGGTCTTCCGCCACCGCCCGT 286
QY 222 CGCGCCCGCATCGGAGACCTTCTTCCCGCGTTCGCGGACCGGACTCTCCCGAC 281
Db 287 CGCGCCCGCATCGGAGACCTTCTTCCCGCGTTCGCGGACCGGACTCTCCCGAC 346
QY 282 CGAGGCGCAGCCCGGAGCGCATCTCACCTCGTCTCGACGACCTCTCGCGGGGT 341
Db 347 CGAGGCGCAGCCCGGAGCGCATCTCGACCTCGTCTCGACGACCTCTCGCGGGGT 406
QY 342 TCAGGCGCCCTGCGTTCGACACATCAAGATCGCGCCATCACGTCGCGCACCGAGTTCGCC 401
Db 407 TGAGGCGCCCTGCGTTCGACACATCAAGATCGCGCCATCACGTCGCGCACCGAGTTCGCC 466
QY 402 GAGGCGCTATCGCCAAATGCTCGCCAAAGGACCGCGGACCAACGAGCGTTCGCTCGG 461
Db 467 GAGGCGCTATCGCCAAATGCTCGCCAAAGGACCGCGGACCAACGAGCGTTCGCTCGG 526
QY 462 ATTCCGGGTCTCCGGCGTCCGAGTCGTCGCGCCCGAGGGCGCGTGTGGCGGACGAGCG 521
Db 527 ATTCCGGGTCT--TGGCTCCGAGTCTGCGCGCCCGAGGGCGCGTGTGGCGGACGAGCG 584
QY 522 -CCGCGAGTCAAGGCCATCGACA-CCGCGCGGCGTCCGCGCGGTTCGCTCCGG 570
Db 585 CCGCGGGGTGAANGCTATGACACCCGTCGNGNCGCGCGGTGCTTCGG 635

RESULT 8
US-10-042-894A-18
; Sequence 18, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1 Inositol Polyphosphate Kinase
; TITLE OF INVENTION: Genes and Uses Thereof
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(519)
; OTHER INFORMATION: n = A, T, C or G
US-10-042-894A-19

Query Match      14.3%; Score 191.6; DB 14; Length 353;
Best Local Similarity 90.2%; Pred. No. 1.6e-41;
Matches 238; Conservative 0; Mismatches 21; Indels 5; Gaps 3;

QY 667 CTCGAAGCGGTGTCGAGGAGAGACTCTGTTCCACTTCTACTCGCGCGTCAATCTCTG 726
```

```
Db      1  CTCAAGCATGGTGGAGGACAGACTCTGTTCCATCTTACTCGGGCTGATCTCTCTG 60
QY      727  GGCTATGATGCTGCTGAGTCGAGGACGAGAGTGGGGTGGGTGACGGTGAAGCTG 786
Db      61  GGCTATGATGCTGCTGAGTCGAGGACGAGAGTGGGGTGGGTGACGGTGAAGCTG 120
QY      787  GTGGACTTTGCCATGTCGAGGACGAGAGTGGGGTGGGTGACGGTGAAGCTG 844
Db      121  GTGGACTTTGCCATGTCGAGGACGAGAGTGGGGTGGGTGACGGTGAAGCTG 180
QY      845  GGCTGCTGCTGAGTCAAGTTC--GTTTCTGACATGTTCCGAGACTCCTCATAG-C 901
Db      181  AGCTCTGCTAGCTGATCAAGTTCGTTCTTGACATGTTTCCAGANACTCCTTAGACGC 240
QY      902  AGCCTTTGGTCTCTTCTTAAGAGA 925
Db      241  AGCCTTTGGTCTCTTCTTAAGAA 264

RESULT 10
US-10-042-894A-11
; Sequence 11, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1 Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; PRIOR FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1195
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (116)...(1048)
US-10-042-894A-11

Query Match      13.7%; Score 184.6; DB 14; Length 1195;
Best Local Similarity 55.8%; Pred. No. 1.6e-39;
Matches 471; Conservative 0; Mismatches 349; Indels 27; Gaps 5;

QY      56  CGCAGCTCCACCGCGGAGACCAAGTCGCGGCCACCGCGCTCCGCCAGCAAGCTG 115
Db      114  CCATGCTCAAGTCCCGGATCATCAAGTCGCGGTCACCGGGAGAGCGGGGAAGCTG 173
QY      116  GCCCGTCATCGAGGCTCGGCTCTTCAAGCGGCTTCAGGCGGGGACCGTGGGG 175
Db      174  GGCACCTGGTGGATGATTCGGCGGCTCTTATAGGCTCTCCAGAGCGATCATCGGGAG 233
QY      176  AGCAGAGTGCCTTCTATAGGCGTCTCCGCCACCGCGCTCCCGCCCGCATCC 235
Db      234  ACAGGAAGTGGCTTTTACGATCATTTATTTCAATACCGAGATCCAGGTCATTC 293
QY      236  GAGACACTTTCTTCCCGGTTCCAGGACGCGACTCTCTCCCGACGAGCGCGCG 295
Db      294  G---CAAAATTTCTTCTGCTTTACGGAACAACTAAT---TGAGGCGTCTGATGAT 347
QY      296  GGGAGCGGCACTCTACCTCGTCTCGACGACTCTCTCGGGGTTTCAGGCGCCCTGG 355
Db      348  CGGGTCTCAACCTCACTCGTCTCGAGGATCTCGTCTGGGTGCGCAGAACCCATCTC 407
QY      356  TCGCAGACATCAAGATCGGGCGCCATCAGTGGGCCACCGAGTTCCGCGGAGCCCTACATCG 415
```

```
Db      408  TCATGACATCAAGACTCGATCCAGAACATGGTATCCGAGGCGCTCTGAGGAGTACATCC 467
QY      416  CCAAGTGCCTCGCAAGGACCGCGGACCAAGAGCGTTCTGCTCGGATTCGCGGTCTCCG 475
Db      468  AANAAGTCTTAGAAGAAGATCGAATAGCACAGCGTTTCATTTGGGTTTATGAGATTCTG 527
QY      476  GGCTCGAGTCTGTCGG---CCCGAGGGCGCGGTGTTGGCGGACGAGCGCGCCGAGTGA 532
Db      528  GCGTAAGGGTATATCAAAAATAGCGAGCTGGATTTTGGCAACCTGAGAAAGAGGTTGTT 587
QY      533  AGGCCATGACACCGCGCGCTCCGCGCTGCTCCGCGCTACGTGTCTATCC----- 585
Db      588  ATAGCTTTAATCGGACCGGTGTCAGGTCTGAGGAAGTTGTTTCTTCCAACTTGT 647
QY      586  -----GTTCCGACAGGGGATGACTGTGCGCTCCCGCGGCGTGTACGGAGGAAAG 640
Db      648  CTCGGGTCCAAATGTGGATCCGGATTGTTTATGCATCAAAAGTTTACTGTACCGGG 707
QY      641  GTGGAGTCTTGTACAGCTGCGGAGCTCAAGCGGTGTTTCGAGGAGCAGACTCTCTTCC 700
Db      708  GTGGAATTTGGCACAAATGCTTCAAGTGAAGGAATGGTTTGAAGTTTCAGACGAATATC 767
QY      701  ACTTCTACTCGCGTCTGATTTCTTCTGGGCTATGATGCTGTCAGTCGAGTCGAGCGGAG 760
Db      768  ACTTCTATTCTTGTTCACATCTATATATGACAGGAGTCTGCTTTGGACGGC---- 823
QY      761  ATGGGGTGGGTGAGCGTGAAGCTGGTGGACTTTTCCCATGTGGCCGAGGTGATGGGG 820
Db      824  --TGTGCACACCGGAAAGTTAACTGGTGGACTTTGACATGTGATGATGGCCACGGCG 881
QY      821  TGAATGACACAACTTCCTGGCGGGCTCTGCTCGCTGATCAAGTTCGTTTCTGACATTC 880
Db      882  TGATCGATCAAACTTCTTGGTGGCTCTGTTCTGTAATCAAGTTTATACGTGACATTG 941
QY      881  TTCGGA 887
Db      942  CTGATGA 948

RESULT 11
US-10-042-894A-9
; Sequence 9, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1 Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; PRIOR FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1105
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (12)...(851)
US-10-042-894A-9

Query Match      13.3%; Score 179; DB 14; Length 1105;
Best Local Similarity 55.6%; Pred. No. 5e-38;
Matches 460; Conservative 0; Mismatches 340; Indels 27; Gaps 5;
```

```
Db 15 CTCAAGATCCCGAGCACCAAGTGGCGGCAAGGCCCAAGGACGAATCTGGGCCCA 74
QY 121 CTCATGAGCGCTCCGGCTCTTCTAAGACCGCTCCAGCGCGGACCGTGGGGAGCAC 180
Db 75 CTCGTGACGATTTGGAAATTCACAGCCCTCCAGACCACAAAGACGACAC 134
QY 181 GAGGTGCGCTTATGAGGCGTT-----CTCGCCCAACCGCGCTCCCGGCCGATC 234
Db 135 CGCGGCTCCACCAACTCTCTTTACACCTCTCTCGCGCGCGCCGCCACGACTACTCC 194
QY 235 CGAGACACCTTCTCCCGCGTTCCAGCGACCGGACTCTCCCAACGAGGCGACGCC 294
Db 195 ATCCGCTCTCTTCTCCCGCGTTCCAGCGACCGGCTCTCTGACGCCCTCCGACGGCTCC 254
QY 295 GGGAGCGCGATCCTCACCTCGCTCCGACGACTCTCTCGCGGGTTTCAGGGCGCCCTGC 354
Db 255 GG---TCCCAACCTCACCTGGTCTCTGGAGGACTCTCTCTCGGCTACTCCAAACCTCC 311
QY 355 GTGCGAGACATCAAGATCGGCGCCATCAGTGGCCACCGAGTTCGCGCGGAGCCCTACATC 414
Db 312 GTCATGACGCTAAAGATCGCTCCAGAACCTGGCACCTGGAGACTCCGAGGACTACATC 371
QY 415 GCCAAGTGCCTCGCCAGGACCGCGGACCAAGAGGTTCTGCTCGGATTCGCGCTCTCC 474
Db 372 TGAAGTGCCTGAAGAAGGACAGAGATCTCTAGCTTGGCTTGGTTTCAGAAATCTCG 431
QY 475 GCGTCCGAGTCTCGCGCCCGGAGGCGCGCTGTGGCGGACGAGCGCGCGGAGTGAAG 534
Db 432 GAGTCAAGGACTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 491
QY 535 GCATGACACCGCGCGCTCCCGCGTCTCGCGGCTCTCGCGGCT---ACGTGTCATCCGTTGCC 591
Db 492 GCCCATGCGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 551
QY 592 GAGCGAGGATGACTGTGCTCGCGCGCGGTTGACGAGGAGAAAGTGGAGTCTTG 651
Db 552 GATCATCATCCCGATTCGCGCTTTCGAAACGAGGCTCTAC-----GGCGCGGTTTG 602
QY 652 TCACAGTCCGCGAGCTCAAGCGTGTTCGAGGACGACTCTGTTCACATTCCTACTCG 711
Db 603 GAGCGCTTCAGAACTCAAGGACTGGTTCGAGGTTTCAGACGCTGATCATTCTATTCT 662
QY 712 GCGTCAATTTCTTGGCTATGATGCTGCTGCACTGCGACGCGGAGATGGGGTGGG 771
Db 663 TGTCTGTTCTTGTGGTGTACGAGAGGATCTAGGAAA-----GGGAAAGCTACCAAC 716
QY 772 GTACAGGTGAAGCTGTGGTACTTTGCCATGTGGCGGAGGTTGATGGGGTGAATGACCA 831
Db 717 CCTCTGGTCAAACTCGTTGACTTTGCACACGTTGGTGGACGGAACGGTGTCTATGATC 776
QY 832 AACTTCTGGCGGGCTCTGCTCGCTGATCAAGTTGTTCTGACAT 878
Db 777 AACTTCTGGGTGGGCTTTGTTCTCTCATCAAGTTCTCTCAAGGATAT 823
```

RESULT 12

```
US-10-424-599-106467
; Sequence 106467, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 106467
; LENGTH: 1345
```

```
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_67156C.1
US-10-424-599-106467
```

Query Match 13.3%; Score 179; DB 12; Length 1345;

Best Local Similarity 55.6%; Pred.No.5.3e-38; Mismatches 340; Indels 27; Gaps 5; Matches 460; Conservative 0;

```
QY 61 CTCACCCCGCGGAGCACCAAGTCCCGCGCCACCGCGCTCCGCCAGCAAGTGGGCCCG 120
Db 220 CTCAGATCCCGGAGCACCAAGTGGCGGCAAGGCCAAGGACGGAATCTCTGGCCCA 279
QY 121 CTCATGACGGTCCCGGCTCTTCTAAGCGCTCCAGCGCGGACCGTGGGGAGGAC 180
Db 280 CTCGTGACGATTTTGGAAATTCACAGCCCTCCAGACCAACAAAGACGACGACCC 339
QY 181 GAGGTGCGCTTCTATGAGGCGTT-----CTCGCCCAACCGCGCTCCCGGCCGATC 234
Db 340 CGCGGCTCCACGAATCTCTTTTACACTCTCTCGCGCGCGCCGCCACGACTACTCC 399
QY 235 CGAGACACCTTCTCCCGGTTTCCACGGCACCGGACTCTCTCCCAACGAGGCGGAGCC 294
Db 400 ATCCGCTCTCTTCTCCCGCTTTCAGGSCACCGCGCTCTCTGACGCTCCGAGGCTCC 459
QY 295 GGGAGCGCGATCTCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 354
Db 460 GG---TCCCAACCTCACCTGGTCTCTGGAGGACTCTCTCTCGGCTACTCTCAACCCCTCC 516
QY 355 GTCGAGACATCAAGATCGGCGCCATCAGTGGCCACCGAGTTCGCGGAGCCCTACATC 414
Db 517 GTCATGAGATGAAGATCGGCTCCAGAACCTGGACCTGGAGACTCCGAGGACTACATC 576
QY 415 GCAAAGTGCCTCGCCAAAGACCGCGGACCAAGAGGTTCTCTCGGATTCGCGGCTCTCC 474
Db 577 TCAAGTGCCTGAAGAGGACAGAGTCTCTAGCTTGGCTTGGTTTCAGAAATCTCG 636
QY 475 GCGTCCGAGTCTGCGGCCCGAGGCGCGCTGTGGCGGACGAGGCGCGCGGAGTGAAG 534
Db 637 GGAGTCAAGGACTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 696
QY 535 GCATGACACCGCGCGCTCCCGCGCTCTCGCGGCT---ACGTGTCATCCGTTGCC 591
Db 697 GCGCATGTTGTTGACTTGTCTCAACAAAGTCTGTTTCTCTAATAATATCAACATGAT 756
QY 592 GACGAGGAGATGAGACTGTGCTCGCTCGCGCGGCTGTACGAGGAAAAGTGGAGTCTTG 651
Db 757 GATCATCATCCCGATTGCGCTTTGCAACGAGGTTCTAC-----GGCGCGGTTTG 807
QY 652 TCACAGCTGGCGAGCTCAAGCGGTTGTCGAGGAGCAGACTCTGTTCACATCTTACTCG 711
Db 808 GAGCGCTTGCAAGAGCTCAAGGACTGGTTCGAGGTTTCAGACGTTGATCATTCTATTCT 867
QY 712 GCGTCAATTTCTGGGCTATGATGCTGTCAGTTCGACGTCGACGAGCGGAGATGGGGTGG 771
Db 868 TGTCTGTTCTTGTGGTGTACGAGGATCTAGGAAA-----GGGAAAGCTACCAAC 921
QY 772 GTGACGTTGAAGCTGGTGGACTTTGCCATGTGGCGGAGGTTGATGGGGTGAATGACCA 831
Db 922 CCTCTGTTCAAACTCGTTGACTTTGCACACGTTGGTGGACGGAACCGGTGTCTATTGATC 981
QY 832 AACTTCTGGCGGCTCTGCTCGCTGATCAAGTTGTTCTGACAT 878
Db 982 AACTTCTGGGTGGGCTTTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1028
```

RESULT 13

```
US-10-042-894A-13
; Sequence 13, Application US/10042894A
; Publication No. US2003009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
```

```
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011alel Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042, 894A
; PRIOR FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1020
; TYPE: DNA
; ORGANISM: Parthenium argentatum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (21)...(908)
; US-10-042-894A-13

Query Match      11.1%; Score 149.8; DB 14; Length 1020;
Best Local Similarity 53.2%; Pred. No. 3.5e-30;
Matches 456; Conservative 0; Mismatches 362; Indels 39; Gaps 5;

QY 61 CTCACCCCGCGGAGCACCAAGTCGGCGGCCACCGCGCTCCCGCAGCAAGCTGGGCGCG 120
Db 24 CTCAAGGCCCCAGATCATCAGGTGCTGGACATGAAGCTGGGCTCGGMAAGCTTGGCCCA 83
QY 121 CTCATCGACGGCTCCGGCTCTTCTACAGCGCTCCAGCGCGCTCAGCGCGCGGAGGCAC 180
Db 84 CTCATTGATGATTGAGCGCGGTTTACAAACCATGTCAGGGTGATACCCGTTGGGTGAGAA 143
QY 181 GAGGTGCGCTTCTATGAGCGGTCTTCGCCCGCAGCGCGCTCCCGCGCGCCATCCGAGAC 240
Db 144 GAAGTAGCCTTTTATGATCATTTCTTCTTAAATAATATTCAGAACACATACG--C 200
QY 241 ACCTTCTTCCCGGTTCCAGGACGCGACTCTCTCCCGCACCAGGCGGACCGCGGGAG 300
Db 201 AAATTCTTCTATATATTATGACCAAAAT-----CATGAAGGCATCCACTGGCTCT 254
QY 301 CCGATCTCTACCTGCTCTCGAGACCTCTCGGGGTTTCAGGGCGCTCGTCCGCA 360
Db 255 GACATCTCTCACATGGTGTTCAGAGATCTTACATGAGCTCATGTCAACCCATCTGTAATG 314
QY 361 GACATCAAGATCGCGGCATCAGGTGGCCACCGAGTTTCGCGGAGCCCTACATCGCCAAG 420
Db 315 GACATCAAAATCGGGTCCAGAACATGGCGCCAGAGCTTCCGAGGGGTACATTCGAAA 374
QY 421 TGGCTCGCCAAAGACCGCGGAGCACGAGGTTCTGTCTCGGATTCGCGGCTCGCGGCTC 480
Db 375 TGGCTTAAAAAAGGATAGGAAAGCACAAAGTATTCATTTGGGATTCAGGATCTCGGGCTG 434
QY 481 CGAGTCTGCGCCCGGAGGCGCGTGTGCGGACG--AGCGCCGAGGTTGAAGGCC 537
Db 435 CAAGTCTATATGATGATGGTTCAGGGTTTTTAAGCCTCATAGAAATTCATCGGTAAA 494
QY 538 ATGACACCCCGCGCTCCCGCGCTGCTCCGGCGCTACGTGTCATCCGTTGCCGAGAG 597
Db 495 ACCGGCCAGCTGATGTTAGACTTCTTATAGAAATTTGTTCTTCTAACCCTCTGCA 554
QY 598 GGGAT-----GGACTGTGCGCTCGCGCGCTAGCGCGGCGGCGGTGACGGA 633
Db 555 GAGATGGAATGCGCACAGGCCTAGGCCGATTTGTTTAGCATCTTTTGTATTAGGT 614
QY 634 GGAAGAGGTGAGTCTTGTACAGCTCGGAGCTCAAGGCTGAGGCTGTCGAGGACAGACT 693
Db 615 GGGCCTAATGGGATTAATAGCTCAACTGATGATGAAATGAAAGACATGGTTGAAAGATCAACA 674
QY 694 CTGTTCACACTTCTACTCGGCGTCAATTTCTCTGGGCTATGATGCTGTCAGTCGACGCA 753
Db 675 ATTACCACTCTATGCTGTTCTTTTGTTCATCTTTGAAAGAGGTTGGTGTAAA 734
```

```
QY 754 GCGGAGATGGGGTGGGTGACGGTGAAGCTGGTGGACTTTGCCCATGTGGCGGAGGT 813
Db 735 GGT--GCTCGGTCAAAACGACAGAGTCAAACTTAATGATTTTGCTCATGTACAGATGGT 791
QY 814 GATGGGTGATTGACCAACAACTTCCCTGGCGGCTCTGCTCGCTGATCAAGTTCGTTTCT 873
Db 792 AATGGTGTATTGATCACAATTTCTTGGGTGGGCTCTGTTCTTTGATAAAGTTCATTTCT 851
QY 874 GACATTGTTCCGGAGAC 890
Db 852 GACATACTTTCGGAGAC 868
```

RESULT 14

```
US-09-770-444-340/c
; Sequence 340, Application US/09770444
; Patent No. US20020023280A1
```

GENERAL INFORMATION:

```
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Kricker, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2027 (PARA-016PRV)
; CURRENT APPLICATION NUMBER: US/09/770,444
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,502
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 340
; LENGTH: 464
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-770-444-340
```

```
Query Match      7.8%; Score 104.6; DB 9; Length 464;
Best Local Similarity 64.2%; Pred. No. 4.2e-18;
Matches 174; Conservative 0; Mismatches 94; Indels 3; Gaps 1;
```

```
QY 618 CGCGCGGTGTACGAGGAAAGTGGAGTCTTGTACAGCTCGCGAGTCAAGGCTG 677
Db 464 CTCGAGTGTATTAGCGGGTTCACCGGGATCTTAACGCGAGTTGCTGGAACCTCAAGACCTG 405
QY 678 GTTCGAGGACGACACTCTGTCACCTTCTACTCGGGTGGGCTCTTCTGGGCTATGATGC 737
Db 404 GTTCGAGAACCAACCCCTCTACCATTTCAACTCTTGTTCGATTTTAAATGCTATAGAA 345
QY 738 TGCTGCAGTCGCA---GCAGGCGGAGATGGGGTGGGGTGACGGTGAAGCTGGTGAATT 794
Db 344 TGAATCACTTAAAGGGAATGATGATGCTAGACCAAGTCAAGTCTGATGGATTT 285
QY 795 TGCCCATGTGGCGAGGATGGGTGATGACCAACTTCTCTGGCGGGCTCTGCTC 854
Db 284 TGCTCATGTTCTTGATGGTAAATGGTGTCAATGACCACTAACTTCTTGGGTGCTCTTGCTC 225
```

QY 855 GCTGATCAAGTTCGTTTCTGACATTTGTCG 885
Db 224 TTTCATAAACTTCATTCGTCAGATCTTCAG 194

RESULT 15
US-10-314-657-1/c
; Sequence 1, Application US/10314657
; Publication No. US20030175888A1
; GENERAL INFORMATION:
; APPLICANT: SHEN, Ben
; APPLICANT: CHENG, Yi-Qiang
; APPLICANT: TANG, Gong-Li
; TITLE OF INVENTION: Discrete Acyltransferases Associated with Type I Polyketide
; TITLE OF INVENTION: Synthases and Methods of Use
; FILE REFERENCE: 054030-0021
; CURRENT APPLICATION NUMBER: US/10/314,657
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US02/08937
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 60/278,935
; PRIOR FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 135638
; TYPE: DNA
; ORGANISM: Streptomyces atroolivaceus
US-10-314-657-1

Query Match 5.3%; Score 71.4; DB 14; Length 135638;
Best Local Similarity 43.9%; Pred. No. 1.4e-08;
Matches 306; Conservative 0; Mismatches 391; Indels 0; Gaps 0;
QY 59 ACCTCCACCCGCGGAGACCAAGTCGCGGCCACCGCGCTCGCCAGCAAGCTGGCC 118
Db 40684 AGCGCGCGTGGCAGAGCTAGCCCCGAGCGCGCGCTGCTCGCCCCCACCCTGGCCC 40625
QY 119 CGCTCATCGAGCGTCCGGGCTCTTCTAAGCCGCTCCAGCCGCGGACCGTGGGAGC 178
Db 40624 GCGCAGGGCGGCTCGGCATCTCGGTACGGCTCTGCCGCGCACCGGGGAGGCACAGC 40565
QY 179 ACGAGGTCGCTTCTATGAGCGTTCTCCGCCACGCGCGTCCCGCGCCGCGATCCGAG 238
Db 40564 GCTACGTGGCATCCCGGCCGACGACAGCTCTACTACAGTACAGCTCGACCCGTCGG 40505
QY 239 ACACCTTTCCCGCGGTTCCAGGCGAGGACCTCTCCACCGAGCGGAGCCCGGG 298
Db 40504 CGGTGAACCTACCTGCTGCCGATCGCGGTGCGGCTGCGGCGACCTGGACACCCCAAGGC 40445
QY 299 AGCGCATCTCTACCTGCTCTCGACACCTCTCGCGGGGTTTCAGGCGCCCTGCGTCG 358
Db 40444 TGGAGCGGGCCCTCGCGCGTCTGCGCCGCGCACAGGCGCTGCGCACCCGCTTCGGA 40385
QY 359 CAGACATCAAGATCGGCGCCATACAGTGGCCACCGAGTTCCCGGAGGCCCTACATCGCA 418
Db 40384 CCGACCCCGAACACGGCGTGTGTCAGTCTGTCGCGACCCCGGAGGTGCGCTCGCAC 40325
QY 419 AGTGCTTCGCAAGGACCGGGACACGAGGTTCTGCTCGGATTCGCGGTCTCCGSG 478
Db 40324 TCGGACGCGCGCGCGCGCGCGCGGAGGTGCGCGACTCTCTGACCGGATCGCCCGGC 40265
QY 479 TCCGAGTCGTGCGGCCCGAGGCGCGCTGTGTGGCGACGAGCGCCCGGAGGTGAAGCCA 538
Db 40264 GCCCTTCGACCTGTTCGCGCGCGCTGTGTGGCGGCAAGCTGTGGCAGCTCACCGAGC 40205
QY 539 TGGACACCGCGCGGTTCGCGCGCTGTCCGCGCTACGTGTATCCGTTGCCACGAGG 598
Db 40204 ACCCGAGGCGCGCGCGCGCTCTCTGCTGCTCACCTGTGCTGCCACCATCTGTCG 40145
QY 599 GGATCGATGTGCGCTCCCGCGGCGGTGTACGGAGGAAAGGTGAGTCTTGTACAGC 658
Db 40144 TGGACGCGTGTGCTCGCGCTGCTCTCGAGAACTACCGAGGGGTACGCGGCCGAGC 40085

QY 659 TCGCGAGCTCAAGCGTGGTTCGAGGAGCAGACTCTGTTCACACTTCTACTCGGCGTCGA 718
Db 40084 TCACCGGGGCGCGCGGAACCTGCCCGAATCGCCGTCAGTACCGGACTTCGCGCGCG 40025
QY 719 TTCTTCTGGGCTATGATGCTGTGCTGCACTGCGAGCAGG 755
Db 40024 CCCATCACGCATGGCTGGCGGTCGCGGTGACGAGGCGCGCAGC 39988

Search completed: March 27, 2004, 06:11:15
Job time : 543 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2004, 20:55:34 ; Search time 20 Seconds
(without alignments)
1389.968 Million cell updates/sec

Title: US-10-042-894A-8

Perfect score: 1521
Sequence: 1 MSDLHPHQVAGHRASAK.....IKFVSDIVPETHPTQPLGPS 289

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	673.5	44.3	286	2	T49876
2	207.5	13.6	323	2	T21569
3	183	12.0	351	2	T17246
4	167.5	11.0	268	2	T50224
5	146.5	9.6	242	2	T27999
6	145	9.5	461	2	TJN0129
7	131	8.6	459	2	S13064
8	123	8.1	355	1	RGBYR3
9	118	7.8	1050	2	S54640
10	116.5	7.7	967	2	T41672
11	114.5	7.5	673	2	S41053
12	110	7.2	394	2	T42512
13	110	7.2	461	2	T42513
14	110	7.2	494	2	T42444
15	109.5	7.2	472	2	S17682
16	109.5	7.2	946	2	T25898
17	103.5	6.8	280	2	T75461
18	98	6.4	1546	2	T75461
19	97.5	6.4	332	2	S73031
20	97.5	6.4	362	2	E86996
21	93.5	6.1	756	2	T70815
22	90.5	6.0	486	2	T25639
23	90.5	6.0	1678	2	T35547
24	90	5.9	1155	2	G87477
25	89	5.9	380	2	C33719
26	89	5.9	4613	2	T17409
27	88.5	5.8	370	1	B43306
28	88.5	5.8	377	1	BOAG58
29	88.5	5.8	377	2	AF3249

30	88.5	5.8	867	2	T14777
31	88	5.8	327	2	D83320
32	87.5	5.8	407	2	T30585
33	87.5	5.8	409	2	T75261
34	87.5	5.8	550	2	T10393
35	87	5.7	363	2	H87464
36	87	5.7	391	2	T08338
37	87	5.7	753	1	D72660
38	86.5	5.7	313	2	D87336
39	86.5	5.7	373	2	T70781
40	86.5	5.7	409	1	BOAG55
41	86.5	5.7	454	2	T47253
42	86.5	5.7	1032	2	T65071
43	86	5.7	605	2	T35277
44	85.5	5.6	319	2	D75349
45	85.5	5.6	337	2	B84335

ALIGNMENTS

RESULT 1

T49876

hypothetical protein T211.80 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000

C:Accession: T49876

R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; L

submitted to the Protein Sequence Database, April 2000

A:Reference number: Z24493

A:Accession: T49876

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-286 <BEV>

A:Cross-references: EMBL:AL163912; GSPDB:GN00063; ATSP:T211.80

A:Experimental source: cultivar Columbia; BAC clone T211

C:Genetics:

A:Gene: ATSP:T211.80

A:Map position: 5

Query Match 44.3%; Score 673.5; DB 2; Length 286;

Best Local Similarity 50.4%; Pred. No. 4.1e-48;

Matches 141; Conservative 42; Mismatches 86; Indels 11; Gaps 7;

Qy	4	LHPPEHOVAGHRASAKLGLIDSGLFYKPLQAGDRGEHEVAYEAFSAHAAPARID	63
Db	3	LKVPFOVAGHIAKDGPGLVDDKGRFFKPLQSDSRGEIEVKFYESFSSNTEVEPIH-	61
Qy	64	TFPFRFHGTLLPTEAQPGEHPHLVLDLILAGFQAPCVADIKIGAITWPPSPPEPIAK	123
Db	62	RYFFVYHGTQAV--EGSDGA--AMVLENLLAEYTKPSVMDVKGSRTWYDPAEEYIQK	117
Qy	124	CLAKDRGTTSVLLGFRVSGVRVWG--PGAVMRTPEPVKAMDTAGRRVLRVYVS--SVA	180
Db	118	CLKKDTGTGTTVSSGFRISGFEVYDHKESFWKPKRLIRGLDVGDAARLTAKFVSSNLS	177
Qy	181	DEGM--DCALAAAYGKGKGVLSQLRELKAWFEOTLFHFYSASILLGYDAAVAAGDGG	238
Db	178	DTGSKPOSFASSVGGSHGILITOLLEKTKWFENQTLYHFNSCSILMYENESILKGNDD	237
Qy	239	GG-VTVKLVDFAHVAEGDGVGDHNFGLGCLSLIKFVSDIV	277
Db	238	DARQVKLVDFAHVLDGNGVIDHNFGLGCLSFNFIREIL	277

RESULT 2

T21569

hypothetical protein F30A10.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Sep-2000

C:Accession: T21569

R:Barlow, K.

submitted to the EMBL Data Library, October 1996

hypothetical prote	
hypothetical prote	
transferase homolo	
cytochrome P450 -	
chitinase (EC 3.2.	
nitrogen regulatio	
conserved hypothet	
probable aldehyde	
riboflavin biosynt	
probable citA prot	
virB10 protein - A	
core protein II (i	
hypothetical prote	
probable transfera	
conserved hypothet	
hypothetical prote	

```

A:Reference number: Z19442
A:Accession: T21569
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-323 <WIL>
A:Cross-references: EMBL:D81072; PIDN:CABO3023.1; GSPDB:GN00019; CESP:F30A10.3
A:Experimental source: clone F30A10
C:Genetics:
A:Gene: CESP.F30A10.3
A:Map position: 1
A:Introns: 9/3; 51/1; 236/2; 291/3
C:Superfamily: Caenorhabditis elegans hypothetical protein F30A10.3

Query Match      13.6%; Score 207.5; DB 2; Length 323;
Best Local Similarity 32.4%; Pred. No. 1.3e-09;
Matches 66; Conservative 36; Mismatches 79; Indels 23; Gaps 7

QY    83  EHPHLVLDDLGLAGFQACVADIKGAIWTFPSPPEPIAKCKORGTTSVLGGFVSQ 142
      ||| :||:::| :||| :||| :||| :||| :||| :||| :||| :|||
Db    128  EPFPFFLLFNVAHYTRPCVDLKLTGTQHGDSDASESKRHRQLMKCHSTSATLGVRVVG 187
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY    143  VRVVGEGAVNR-TERPEVKAMDTAGRVRRVLRYSVSSVADEGMDCALAAAYCGKGCVLS 201
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db    188  MOLYEATKSYSVVEKQEGRRIDAAAGFYGVYKRFIKC-----CGSRRAAIRIQ--KLS 238
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY    202  QUREIKAMEEEQTLEPHFYSASTLLGYDAAAAGAAGDGGVTVKLVDFAHVAEGDGI--- 258
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db    239  KURLSLAEFG---YRFSSSLITAFDAEADSSD-DADVLCIIDPAH-STFSGGFEDL 293
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY    259  -----DHNFLGLCLSLIRFVSIV 277
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

Db    294  AYSGADEGCULGLDSTIVEAMEPIV 317
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

```

RESULT 3
T17246
hypothetical protein DXFPz586M0617.1 - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Sep-2000
C/Accession: T17246
R/Koehler, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A/Reference number: Z18722
A/Accession: T17246
A/Status: Preliminary
A/Molecule type: mRNA
A/Residues: 1-351 <EO>
A/Cross-references: EMBL:AL117458
A/Experimental source: adult uterus; clone DXFPz586M0617
C/Genetics:
A/Note: DXFPz586M0617.1
C/Superfamily: Caenorhabditis elegans hypodermis protein F30A10.3

	Query Match	12.0%	Score 183;	DB 2;	Length 351;
	Best Local Similarity	25.6%;	Pred. No. 1.5e-07;		
	Matches	62;	Conservative	42;	Mismatches 68; Indels 70; Gaps 7;
Qy	88	LVLDDLLAGFOACPADIKIGAITWPPSSPRPYIAKCIACDKRGTTSVLLGPRVSGVRVVG	147		
	:::::	:::::	:::::	:::::	:::::
Dd	129	ILLENLTSRYEVPCLDUMKGTROHGDASEKAANIRKCQQSTSAIVGRVCQMUY-	187		
	:::::	:::::	:::::	:::::	:::::
Qy	148	PEGAVWRTERPEVKAMDTAGVRRVL---RRYVSVADEGMDCALAAAYVGGK----	G 197		
	:::::	:::::	:::::	:::::	:::::
Dd	188	-----QAQSGGLMFMNKYHGKXLSVQGFKREALFQFHNGRYLRRELLG	230		
	:::::	:::::	:::::	:::::	:::::
Qy	198	GVLSQLRELKAWFEQTLFPHYSAIILLGYD-----AAAAAAG	235		
	:::::	:::::	:::::	:::::	:::::
Dd	231	PVLKKLTAKLVLRQESYRFYSLLLVIYDGKRPERVWLDSDAEDLEDLESSEADESAG	290		
	:::::	:::::	:::::	:::::	:::::
Qy	236	G-----DGGGVTVKLVDFAH-----VAEGDGVIDHNFLGGICSLIKFVSIVP	278		
	:::::	:::::	:::::	:::::	:::::
b	291	AVAKPIGASVVDMTDFATTCCTCYGDTVMVHEGO---DAGVFEGLOSHDIVITISE	347		
	:::::	:::::	:::::	:::::	:::::

```

QY      279 ET 280
      |
Db      348 ES 349

RESULT 4
T50224      T50224
C;Species: Schizosaccharomyces pombe
C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jun-2000
C;Accession: T50224
R;Zimmermann, W.; Wambutt, R.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, December 1999
A;Reference number: Z25047
A;Accession: T50224
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-268 <ZIM>
A;Cross-references: EMBL:AL135751; FIDN:CAB63791.1; GSPDB:GN00066; SPDB:SPAC607.04
A;Experimental source: strain 972h(-); cosmid c607
C;Genetics:
A;Gene: SPDB:SPAC607.04
A;Map position: 1

```

Query Match	11.0%;	Score	167.5;	DB 2;	Length	268;			
Best local Similarity	28.7%;	Pred. No. 2e-06;							
Matches	54;	Conservative	29;	Mismatches	88;	Indels	17;	Gaps	5
Qy	88	LVLDDLAGFQAPCVADIKIGAIWFPSSPPFYIAKCIADKRGTTSVLLGFRVSGVRVVG	147						
Db	83	ILLENILYQMETPCVMDIKLGKQLWADDAPLEKRRLLDAVSRSTTSGSLGFRITGI----	138						
Qy	148	PEGAVW--RTERPEVKAMDTAGVRRVLRYSVSADEGMDCCALAAVYGGKGVLSQLREL	206						
Db	139	---LSWDRITNNITYIKRSTAWGKTTTDSDWEGLENDFFVSCSLSQ-----KARLVESFNL	190						
Qy	207	KAWPE---EQLTFHFYSAILLGDDAAAV--AAGDGGGVTVKLVDFAHVAEGDGVIDHNF	262						
Db	191	LKLFVQLSSEYIELKSSILFTVYSSSLNFTYHCEGNNVVKLLDLAHSRWTNTIDHNT	250						
Qy	263	IGGLCSLI	270						
Db	251	LTGVKNLT	258						

RESULT 5

T27999
hypothetical protein ZK795.1 - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C/Accession: T27999
R/Percy, C.
submitted to the EMBL Data Library, December 1996
A/Reference number: Z20453
A/Accession: T27999
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-242 <WLL>
A/Cross-references: EMBL:Z83246; PIDN:CA805842.1; GSFDB:GN00022; CESP:ZK795.1
A/Experimental source: clone ZK795
C/Genetics:
A/Gene: CESP:ZK795.1
A/Map position: 4
A/Introns: 36/1; 59/3; 88/2; 116/1
A/Superfamily: *Caenorhabditis elegans* hypothetical protein ZK795.1

	Query Match	9.6%	Score 146.5	DB 2	Length 242
Best Local Similarity	27.8%	Pred. No. 9.5e-05			
Matches	45	Conservative 22	Mismatches 64	Indels 31	Gaps 5
Qy	10	QVAGHRAASAKLGP-----LIDGSGFYFQAGDRGEHEVAFYEAFSAHAAPVA-----	59		
Db	19	OTAGHP-SVINGKREIGLLIKIPGSRRIILKPKDASRGEKEVALYELLRSCTTSPTTPE	78		

QY 60 -----RIRD-----TFRFHGTRLLPTEAQGPHPHLVLDLLAGFOAPCVA 103
| : | | : | : | : | : | : | : | : | :
DB 79 STSDELKRVREDIEMCKLTAFYGIOTFVDGODRE---FLANEDVTIGYORPAIL 135

QY 104 DIKIGATTWPSSPEPIAKCLAKDRGTTSVLLGFRVSGVR 145
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
pB 136 DLKMGVTVDPIAKPEKIEK--ERIKYPPOAKGMIIILGYRV 175

RESULT 6

JN0129
1D-myo-inositol-trisphosphate 3-kinase (EC 2.7.1.127) A - human
N;Alternate names: inositol-1,4,5-trisphosphate 3-kinase A
C;Species: Homo sapiens (man)
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 21-Jul-2000
C;Accession: JN0129; S13559
R;Takazawa, K.; Perret, J.; Dumont, J.E.; Erneux, C.
Biochem. Biophys. Res. Commun. 174, 529-535, 1991
A;Title: Molecular cloning and expression of a human brain inositol 1,4,5-trisphosphate
A;Reference number: JN0129; MUID:91128380; PMID:1847047
A;Accession: JN0129
A;Molecule type: mRNA
A;Residues: 1-461 <TAK1>
A;Cross-references: EMBL:X54938; NID:G32104; PID:G32105
R;Takazawa, K.; Perret, J.; Dumont, J.E.; Erneux, C.
Nucleic Acids Res. 18, 7141, 1990
A;Title: Human brain inositol 1,4,5-trisphosphate 3-kinase cDNA sequence.
A;Reference number: S13559; MUID:91088302; PMID:2175886
A;Accession: S13559
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-461 <TKA2>
A;Cross-references: EMBL:X54938; NID:G32104; PIDN:CAA38700.1; PID:G32105

[illegible]

DB 346 IPGNEDGYLGIDNLVILDEL-PE 370

RESULT 13

T42513

1D-myo-inositol-trisphosphate 3-kinase (EC 2.7.1.127) isoform 3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
C:Accession: T42513
R:Cladinin, T.R.; DeModena, J.A.; Sternberg, P.W.
Cell 92, 523-533, 1998
A>Title: Inositol trisphosphate mediates a RAS-independent response to LET-23 receptor

A:Reference number: Z22166; MUID:98150857; PMID:9491893
A:Accession: T42513
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-461 <CLA>
A:Cross-references: EMBL:AF045611; NID:g2898161; PIDN:AAC38962.1; PID:g2898162
C:Genetics:
A:Gene: LFE-2
C:Keywords: phosphoprotein; phosphotransferase

Query Match 7.2%; Score 110; DB 2; Length 494;
Best Local Similarity 21.8%; Pred. No. 0.23;
Matches 71; Conservative 46; Mismatches 99; Indels 110; Gaps 18;

QY 10 QVAGHRASAKGLPLIDGSLFYKPLQAGDRGHEVAFAEFAHAAVPAIRDTFFPRF 69
DB 166 QLSGHEGIVPATP-----HTLYKK-QCANCGBGRA--YKNISKDPAL-----DGTFKY 212
QY 70 HGTRLLPTEAQGPGEHPHLVLDLLAGFQAP---CVADIKIGAITWPPS-----115
DB 213 Y-----DELEKNEH-FIEIEDLQQFHDPTKTAIMDKIGTRTFLESEVSNTKKRADL 265
QY 116 -----SPEPIAKCLA-----KDRGTSVLLGFRVSGVRVVGPEGA-----151
DB 266 YERKWAINDPEETEERKCGAITKLRVQFRRESSTAQIGFRIEAAKRL--EGALEKNF 323
QY 152 --VWRTPEVKAMDTAGVRRVRLRYVSSVADEGMDCALAAAVYGGKGVLSQLRELKAW 209
DB 324 KKVRTVEDVTTFMDFGQTOR-----SRVROQ-----LIERLKSWMKA 361
QY 210 FEEQTLPHFY---SASILLGYDAAAVAGGDDGGVTVKLVDFAH---VAEG-----254
DB 362 IEHSSFFNSHEVVGSSILIVDFTEKVGCM-----MIDFAKSSPVNGRTLNRHTTW 412
QY 255 -DGVIDHNFGLGCLSLIKFVSDIVPE 279
DB 413 IPGNNEEDGYLIGIDNLVKILEEL-PE 437

RESULT 14
T42444
1D-myo-inositol-trisphosphate 3-kinase (EC 2.7.1.127) 1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
C:Accession: T42444
R:Clandinin, T.R.; DeModena, J.A.; Sternberg, P.W.
Cell 92, 523-533, 1998
A:Title: Inositol trisphosphate mediates a RAS-independent response to LET-23 receptor
A:Reference number: Z22166; MUID:98150857; PMID:9491893
A:Accession: T42444
A:Molecule type: mRNA
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Residues: 1-494 <CLA>
A:Cross-references: EMBL:AF045611; NID:g2898157; PIDN:AAC38960.1; PID:g2898158
C:Genetics:
A:Note: LFE-2
C:Keywords: phosphotransferase

Query Match 7.2%; Score 110; DB 2; Length 494;
Best Local Similarity 21.8%; Pred. No. 0.23;
Matches 71; Conservative 46; Mismatches 99; Indels 110; Gaps 18;

QY 10 QVAGHRASAKGLPLIDGSLFYKPLQAGDRGHEVAFAEFAHAAVPAIRDTFFPRF 69
DB 199 QLSGHEGIVPATP-----HTLYKK-QCANCGBGRA--YKNISKDPAL-----DGTFKY 245
QY 70 HGTRLLPTEAQGPGEHPHLVLDLLAGFQAP---CVADIKIGAITWPPS-----115
DB 246 Y-----DELEKNEH-FIEIEDLQQFHDPTKTAIMDKIGTRTFLESEVSNTKKRADL 298
QY 116 -----SPEPIAKCLA-----KDRGTSVLLGFRVSGVRVVGPEGA-----151
DB 299 YERKWAINDPEETEERKCGAITKLRVQFRRESSTAQIGFRIEAAKRL--EGALEKNF 356

QY 152 --VWRTPEVKAMDTAGVRRVRLRYVSSVADEGMDCALAAAVYGGKGVLSQLRELKAW 209
DB 357 KKVRTVEDVTTFMDFGQTOR-----SRVROQ-----LIERLKSWMKA 394
QY 210 FEEQTLPHFY---SASILLGYDAAAVAGGDDGGVTVKLVDFAH---VAEG-----254
DB 395 IEHSSFFNSHEVVGSSILIVDFTEKVGCM-----MIDFAKSSPVNGRTLNRHTTW 445
QY 255 -DGVIDHNFGLGCLSLIKFVSDIVPE 279
DB 446 IPGNNEEDGYLIGIDNLVKILEEL-PE 470

RESULT 15
S17682

1D-myo-inositol-trisphosphate 3-kinase (EC 2.7.1.127) B - human
N:Alternate names: inositol-1,4,5-trisphosphate 3-kinase B
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 21-Jul-2000
C:Accession: S17682
R:Takazawa, K.; Perret, J.; Dumont, J.E.; Erneux, C.
Biochem. J. 278, 883-886, 1991
A:Title: Molecular cloning and expression of a new putative inositol 1,4,5-trisphosphat
A:Reference number: S17545; MUID:91378954; PMID:1654894
A:Accession: S17682
A:Molecule type: mRNA
A:Residues: 1-472 <PAK>
A:Cross-references: EMBL:X57206; NID:g33990; PIDN:CAA40491.1; PID:g33991
C:Genetics:
A:Gene: GDB:ITPKB
A:Cross-references: GDB:128973; OMIM:147522
A:Map position: lq41-lq43
C:Keywords: phosphotransferase

Query Match 7.2%; Score 109.5; DB 2; Length 472;
Best Local Similarity 20.3%; Pred. No. 0.24;
Matches 55; Conservative 40; Mismatches 85; Indels 91; Gaps 13;

QY 65 FPFRTGTRLLPTEAQGPGEHPHLVLDLLAGFQAPCVADIKIGAITW-----112
DB 237 FVPAYHG-----DVVKDGERVNO--MDDLIAIDFSPCWMCKMGIRTYLEELTKARKKP 289
QY 113 -----PPSP--EPYIAKCLAKORG-----TTSVLLGFRVSGVRVVGPEG 150
DB 290 SLRKMVQRMIEVDPEAPTEEERKQAVTKPRVQWRETISTATLGFRIEGIK--KEDG 347
QY 151 AVWRTPEVKAMDTAGVRRVRLRYVSSVADEGMDCALAAAVYGGKGVLSQLREL 206
DB 348 TV---NRDFKTKTREQVTEAPREFYK-----GNHNILIAVRDLKAI 387
QY 207 KAWFEEQTLPHFY---SASILLGYDAAAVAGGDDGGVTVKLVDFAH---VAEG-----254
DB 388 RTTLEVSPPFKCHEVIGSSLLFIHDKKEQA-----KVMIDFGKTTPLPEGQTLQHD 439
QY 255 ----DGVIDHNFGLGCLSLIKFVSDIVPE 281
DB 440 VPMQEGNREDGYLSGLNVLVILTEMSQDAP 470

Search completed: March 24, 2004, 20:59:40
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 27, 2004, 06:02:11 ; Search time 96 Seconds
(without alignments)
1670.633 Million cell updates/sec

Title: US-10-042-894A-8

Perfect score: 1521

Sequence: 1 MSDLHPBHQVAGHRASAK.....IKFVSIVPETHPTQPLGPS 289

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 27475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=framer_p2n.model -DEV=xlh
-Q=/cgn2_1/uspto.spool/US10042894/runat_24032004_145049_16992/app_query.fasta_1.455
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10042894 -CGN_1_69 @runat_24032004_145049_16992 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:

1: /cgn2_6/prodata/2/ina/5A COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B COMB.seq:*
5: /cgn2_6/prodata/2/ina/6C COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	193	12.7	1341	4	US-09-023-655-809
2	189	12.4	1748	4	US-09-620-312D-712
3	113.5	7.5	1899	4	US-09-252-991A-6525
C 4	107.5	7.1	1984	4	US-09-252-991A-12872
C 5	107.5	7.1	4078	4	US-09-016-434-1132
C 6	106.5	7.0	6000	1	US-08-348-006B-6
C 7	106.5	7.0	6000	2	US-08-800-825A-6
C 8	106.5	7.0	6000	3	US-09-358-657-6
C 9	106.5	7.0	6000	5	FCI-US94-10166-6
C 10	106	7.0	4411529	3	US-09-103-840A-1
C 11	105.5	6.9	36412	4	US-08-311-731A-132
12	104	6.8	1917	2	US-08-637-899-2

13	103.5	6.8	1896	4	US-09-252-991A-11331	Sequence 11331, A
C 14	100.5	6.6	960	4	US-09-252-991A-12565	Sequence 12565, A
C 15	100.5	6.6	2514	4	US-09-144-914-3	Sequence 3, Appli
C 16	100.5	6.6	3446	4	US-09-620-312D-653	Sequence 653, App
17	100	6.6	1425	4	US-09-489-039A-1509	Sequence 1509, Ap
C 18	100	6.6	1491	4	US-09-252-991A-6021	Sequence 6021, Ap
C 19	100	6.6	2016	4	US-09-489-039A-1492	Sequence 1492, Ap
C 20	99	6.5	1470	4	US-09-489-039A-6625	Sequence 6625, Ap
C 21	98	6.4	4403765	3	US-09-103-840A-2	Sequence 2, Appli
C 22	98	6.4	4411529	3	US-09-103-840A-1	Sequence 1, Appli
C 23	96.5	6.3	1509	4	US-09-489-039A-6892	Sequence 6892, Ap
24	96.5	6.3	4403765	3	US-09-103-840A-2	Sequence 2, Appli
C 25	96	6.3	1230	4	US-09-252-991A-14780	Sequence 14780, A
C 26	96	6.3	1560	2	US-08-709-979A-2	Sequence 2, Appli
C 27	96	6.3	23673	4	US-09-773-816-1	Sequence 1, Appli
C 28	95	6.2	813	4	US-09-489-039A-5090	Sequence 5090, Ap
C 29	95	6.2	1952	4	US-09-566-921-24	Sequence 24, Appli
C 30	95	6.2	2154	4	US-09-488-856A-3	Sequence 3, Appli
C 31	95	6.2	10095	3	US-08-822-586-45	Sequence 45, Appli
32	94.5	6.2	1227	4	US-09-252-991A-15822	Sequence 15822, A
C 33	94.5	6.2	1353	4	US-09-252-991A-15870	Sequence 15870, A
C 34	94.5	6.2	1719	4	US-09-252-991A-15897	Sequence 15897, A
C 35	94	6.2	1257	4	US-09-489-039A-6981	Sequence 6981, Ap
C 36	94	6.2	3765	4	US-09-266-965-9	Sequence 9, Appli
C 37	94	6.2	53500	4	US-09-266-965-76	Sequence 76, Appli
C 38	93.5	6.1	1155	4	US-09-252-991A-14787	Sequence 14787, A
C 39	93.5	6.1	2082	4	US-09-252-991A-2596	Sequence 2596, Ap
C 40	93.5	6.1	2127	4	US-09-818-780-75	Sequence 75, Appli
C 41	93.5	6.1	2388	4	US-09-252-991A-2514	Sequence 2514, Ap
C 42	93	6.1	2401	3	US-09-320-878-20	Sequence 20, Appli
C 43	93	6.1	2401	4	US-09-141-908-9	Sequence 9, Appli
C 44	93	6.1	2401	4	US-09-657-440-20	Sequence 20, Appli
C 45	93	6.1	3147	2	US-08-781-802-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1

US-09-023-655-809
; Sequence 809, Application US/09023655
; Patent No. 6607819
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:

```

/ TELEPHONE: (650) 855-0555
/ TELEFAX: (650) 845-4166
/ INFORMATION FOR SEQ ID NO: 809:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1341 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: BRAITUT01
/ CLONE: 753522
US-09-023-655-809

Alignment Scores:
Pred. No.: 4,75e-12 Length: 1341
Score: 193.00 Matches: 88
Percent Similarity: 40.3% Conservatives: 48
Best Local Similarity: 26.11% Mismatches: 110
Query Match: 12.69% Indels: 94
DB: 4 Gaps: 10

US-10-042-894A-8 (1-289) x US-09-023-655-809 (1-1341)
QY 8 GluHisGlnValAlaGlyHisArgAlaSerAlaSerLys-LeuGlyProLeuIleAspG1 27
Db 176 AAACATCATGCTCTAGGAACAGAAAGAAACCCCTAAGGAGCTGGGTGCGTCCACAGATAA 235
QY 27 ySerGlyLeuPheTyLysPro----- 34
Db 236 AAGAGAGAAATGAGAGCCATAGTAGAGAGAAATTTGAGTGGCTAAAGAAATCTG 295
QY 35 -----LeuGlnAlaGlyAspArgGlyGluHisGluValAlaPheTyrGluAlaPh 51
Db 296 AAGTCTTGTACTACACTGTAGAGAGAGAGGGAATATAAG-----TT 337
QY 51 eSerAla-----HisAlaAlaValProAlaArgIleArgAspTh 64
Db 338 CCCAGCTTAACACTATAACCCCTGGAGCATGAATGTCCAGCAGCAACAGTTACAGAGAA 397
QY 64 rPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGluPr 84
Db 398 T-----GAGGAGAAATGCAAGCATCGGAAC 424
QY 84 o-HisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaProCysValAlaA 104
Db 425 AGTACAAATTTATCTTACTGGAACCTGACTTCCCGCTATGAGTGCCTTGTGCTCTTG 484
QY 104 spLleLysIleGlyAlaIleThrTrpProSerProGluProTyrlleAlaLysC 124
Db 485 ACCTCAAGTGGGCAACAGCAACATGGTGTATGATGCTTCAGAGAGAGAGCCCAACC 544
QY 124 ySLeuAlaLysAspArgGlyThrSerValLeuLeuGlyPheArgValSerGlyValA 144
Db 545 AGATCCGAAATCTCAGCAGAGACATCTGAGTCATTTGGTGTGCTGTGTGGCATGC 604
QY 144 rgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluValLysAlaMetA 164
Db 605 AGGTGTAC-----CAAGCAGGCGAG-TGGCAGCTCATGTTTCATGAACA-AGTACCATTG 656
QY 164 spThrAlaGlyValArgArgValLeuArgArgTyrValSerSerValAlaAspGluGlyM 184
Db 657 ACGAAGCTATCGGTGAGGCTTCAAGAGGACCTTTTCCAGTCTTCCACATGGGCG 716
QY 184 etAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyValLeuSerGlnLeuA 204
Db 717 GTACTCGCGCGTG-----AACTCCTG-GGCCCTGTGCTCAAGAAGCTGA 760
QY 204 rgGluLeuLysAlaTrpPheGluGlnThrLeuPheHisPheTyLysSerAlaSerIleL 224
Db 761 CTGAGCTCAAGGCGAGTGTGGAGCGACAGGAGTCTTACCCTTCTACTCAAGTCCCTGC 820
QY 224 euLeuGlyTyLysPhe----- 228

```

```

Db 821 TGGTCATTATGATGGCAAGGAGCGGCCGAGTGGTCTCTGGACTCAGATGCTGAGGATT 880
QY 229 -----AlaAlaValAlaAlaGlyGly----- 236
Db 881 TGGAGGACCTGTCTCAGAGGAATCAGCTGATGAGTCTGCTGCTGCTATGCTTACAAACCA 940
QY 237 --AspGlyGlyGlyValThrValLysLeuValAspPheAlaHis----- 250
Db 941 TCGGCGCCAGCTCTGTAGATGTGCGATGATGCGATTGTGCACTTTGACACACACCCTGCAGGCTGT 1000
QY 251 -----ValAlaGluGlyAspGlyValIleAspHisAsnPheLeuGlyG 265
Db 1001 ATGCGGAGGACACCGTGTGATGATGAGGCGCAG-----GATGCTGCTATATCTTCG 1051
QY 265 yLeuOysSerLeuIleLysPheValSerAspIleValProGluThr 280
Db 1052 GGCTCCAGAGCGCTGATACATTTGTTCACAGAGATAAGTGAAGAGAGT 1098

RESULT 2
US-09-620-312D-712
; Sequence 712, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 712
; LENGTH: 1748
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (187)..(1467)
US-09-620-312D-712

Alignment Scores:
Pred. No.: 1.99e-11 Length: 1748
Score: 189.00 Matches: 71
Percent Similarity: 45.27% Conservatives: 39
Best Local Similarity: 29.22% Mismatches: 83
Query Match: 12.43% Indels: 53
DB: 4 Gaps: 6

US-10-042-894A-8 (1-289) x US-09-620-312D-712 (1-1748)
QY 78 GluAlaGlnProGlyGluPro-HisProHisLeuValLeuAspLeuLeuAlaGlyPh 97
Db 765 GAATGCAAGCATCGGAACAGTACAAATTTATCTTACTGGAAACCTGACTTCCCGCTA 824

```

QY 97 eGlnAlaProCysValAlaAspIleIysIleGlyAlaIleThrTrpProSerSerPr 117
 Db 825 TGAGGTGCCTTGTCTTCACTCAAGATGGGCACACGCAACATGCTGATGCTTC 884
 QY 117 oGluProTyIleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuG 137
 Db 885 AGAGGAGAGGAGCCACACAGATCCGAAATGTGACGAGGACATCTGCAGTCATGG 944
 QY 137 yPheArgValSerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluAr 157
 Db 945 TGTGCGTGTGTGTCATGAGTGTATC-----CAAGCAGGCGAG-TGGCGAGCTCATGT 997
 QY 157 gProGluValLysAlaMetAspThrAlaGlyValArgArgValLeuArgTrpValse 177
 Db 998 TCATGAACA-AGTACCATGACGAGAACTATCGGTGACGGCTTCAAGGAGGACATTTC 1056
 QY 177 rSerValAlaAspGluGlyMetAspCysAlaLeuAlaAlaValTyGlyGlyLysG 197
 Db 1057 CAGTCTTCCACATGGGCGGTACTTCGCGGTG-----AACTCCTG-GG 1100
 QY 197 yGlyValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHi 217
 Db 1101 CCCTGTGCTCAAGAGAGTGAAGTGAAGTCAAGGAGTGTGTGAGGCGACAGGAGTCTTACCG 1160
 QY 217 sPheTySerAlaSerIleLeuLeuGlyTyTrAsp----- 228
 Db 1161 CTTCTACTCAAGCTCCCTGCTGGTCNTTATGATGGCAGGAGCGGCCGAGTGTCT 1220
 QY 229 -----AlaAlaAlaValAlaAlaG 235
 Db 1221 GGACTCAGATCTGAGGATTTGGAGGACCTGTCAGAGGAATCAGTGTGATGCTGCTGG 1280
 QY 235 yGly-----AspGlyGlyValThrValLysLeuValAspPheAl 249
 Db 1281 TGCCTATGCTTACAAACCCATCGCGCCAGCTCTGTAGATGTGCGCATGTGCTGTC 1340
 QY 249 aHis-----ValAlaGluGlyAspGlyValI 258
 Db 1341 ACACACACCTGCAGGCTGTATGGGAGGACACCGTGTGTCATGAGGCGCAG- 1392
 QY 258 eAspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValPr 278
 Db 1393 -GATGCTGCTATATCTTCGGGCTCCAGAGCTGATGACATTTGTACAGAGATAAGTGA 1451
 QY 278 oGluThr 280
 Db 1452 GGAGAGT 1458

RESULT 3

US-09-252-991A-6525/c
 ; Sequence 6525, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 6525
 ; LENGTH: 1899
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-6525

Alignment Scores:

Pred. No.: 0.00696 Length: 1899

Score: 113.50 Matches: 93
 Percent Similarity: 33.51% Conservative: 35
 Best Local Similarity: 24.35% Mismatches: 115
 Query Match: 7.46% Indels: 139
 DB: 4 Gaps: 20
 US-10-042-894A-8 (1-289) x US-09-252-991A-6525 (1-1899)
 QY 12 AlaGlyHisArgAlaSerAlaSerLysLeuGlyProLeuIleAspGlySerGlyLeuPhe 31
 Db 1515 GCGCGCACCGCTACTTGGTGGCGTGTAGGCGCT-----CATCGCGGCTGGC 1465
 QY 32 TyLProLeuGlnAlaGlyAsp-----ArgGlyGluHisGluVal 45
 Db 1464 GAACAGCCCGAAGACGCTGAGACATTTGACCAGTGGCCGCTTCCGAGGCGCTTCAGGTG 1405
 QY 46 AlaPheTyGluAlaPheSerAlaHisAlaAlaValProAlaArgIleArgAspThrPhe 65
 Db 1404 CGGAAGAACGCTTGGTCCCTTCCACCGCCACAGAGTTGATGTCATGATCCACTC 1345
 QY 66 PheProArg-----PheHisGlyThrArgLeuLeuProThrGluAlaGlnPro 81
 Db 1344 GTATTCGAGTACTGCTGCTTCCACGCTACCGCGTGGCGAGCGCCGCGTGTGTAA 1285
 QY 82 GlyGluProHisPro-----HisLeuVal 89
 Db 1284 AATCAGGTTCCCGCGCATGTTGCTGGCGCCTTGTGCGCCAGCGCTGCACCTGCTC 1225
 QY 90 -----LeuAspAsp-----LeuLeu 94
 Db 1224 GCGTCCGCGCAGCTCGACACCGCGGTGGAGACCGCTACCCGCTGCTGGCGAGCAATTG 1165
 QY 95 AlaGlyPheGlnAlaProCysValAlaAspIleLysIleGlyAlaIleThrTrpProPro 114
 Db 1164 GCGGTTTCTCGAGAGCTGCGCATTTGACGTGCGCCAGGCGCAGGTCACACCTGCTC 1105
 QY 115 Ser--SerProGluProTyTrIleAlaLysCysLeuAlaLysAspArg----- 129
 Db 1104 GCCCAGCTCCACGCGCAGGCGCGACCGATGCGCGGCGCTGCGCGGAGTGGCGGCGAC 1045
 QY 130 -----GlyThrThrSer-----ValLeuLeuGlyPheArgVal 141
 Db 1044 TTTGTTTTCGAACGACTTCATGCGGGGTACCCCTTCGCGCTGTTGGCAGGCTTCGTTGCTC 985
 QY 141 erGly-----ValArgValValGly----- 147
 Db 984 AGGCGCGGAGCGGTGCTGCTGAGTGTGTAGGCGGAGCGCTCGAACTGCGGCTCAGG 925
 QY 148 -----ProGluGlyAlaValTrpArgThrGluArgProG 159
 Db 924 GCACGGAAGCGCAGGTGAATCCCGGCCACACAGGTCGAGTTGCGGCGCTGACCGGATGC 865
 QY 159 luValLysAlaMetAspThr-----AlaGlyValArgArgValLeuArg 174
 Db 864 AGGACACGCTCTTTCGAACCGCCACCGCTCCAGACGCTGCGCTCCAGTTGCGTGGAGG 805
 QY 174 rGlyrValSerSerValAlaAspGluGlyMetAspCysAlaLeuAla-----AlaAlaV 192
 Db 804 TATTCGTTGTAGCGTTCCTCGACCGCGGCGCTTGAACCTCGAGACTGAGCAGTTTCGCGGCGC 745
 QY 192 alTyGly----- 194
 Db 744 TTCATCAGCTTCAGCGCATCGAGGACGTGAGCGATCTCGACTCGCATCATGTAGACCATC 685
 QY 195 -----GlyLysGlyGlyValLeuSerGlnLeuArgLulLeuLysAlaTrpPheGluG 212
 Db 684 GAGTTGTGCGGAGCGCGGTGTTCCGTCCCAT-----CAGG 649
 QY 212 luGlnThrLeuPheHisPheTySerAlaSerIleLeuLeuGlyTyTrAspAla----- 229
 Db 648 AAGACAGGTTGGGAAAGCCGCGGTGTGTAGTGTGCTTGTAGGCTTCGCGGCGCTTGTGTC 589
 QY 230 -----AlaAlaValAla-----AlaGlyGlyAsp----- 237

		:::		
Db	588	CAGCTGTCACAGGTTCGGCGCGTTCGCACCGAAGACCCTCCGCGGGGATAGGATCG	529	
	238	--GlyGly-----GlyValThrValLysLeuValAspPheAlaHisValLa-----	252	
	528	TTGCGGTGAACCGGTACCGAAGATGATCGCGTCGACCTCGCGTTCCGGCGGTTGCCG	469	
	253	-----GlucGlyAspGlyValIle-----AspHisLenPheLeuGlyG	265	
Db	468	TCGACGATTCCATTGGCGGTGACGCGCGGGATCGCTCGGTGATCACCGTGAGTTGGCG	409	
	265	ly	265	
Qy	408	GC	407	
b				

RESULT 4

```

RES001.
US-09-252-991A-12872/c
; Sequence 12872, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12872
; LENGTH: 984
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-12872

```

```

Alignment Scores:
Pred. No.:      0.0123      Length:      984
Score:          107.50      Matches:      91
Percent Similarity: 35.86%      Conservative: 32
Best Local Similarity: 26.53%      Mismatches: 117
Query Match:      7.07%      Indels:      105
DB:              4          Gaps:      15

US-10-042-894A-8 (1-289) x US-09-252-991A-12872 (1-984)

QY 8 GluHisGlnValAlaGlyHisArgAlaSer-----AlaSerIysLeuGlyPro 23
Db 968 GAATACGCAATGGCCGGAGCAGCTCTAGTCGACCGGCCATCGGAATACTCGCGGGCG 909
QY 24 LeuIleAspGlySerGlyLeuPheTyrIysProLeuGlnAlaGlyAspArgGlyGlu--- 42
Db 908 TCCGGCGGACGGCGAGGCACAGTTT-----ACGGTCAGCGCTGGATTTCCGCGTGC 855
QY 43 -----HisGluValAlaPheTyrGluAlaPhe----- 51
Db 854 CCAGCGTGTAAAGCGAGATTGAGAGACAGCGGCGCTCTCTCGGCGCTTCGAGGCTGA 795
QY 52 -----SerAlaHisAlaValPro-AlaArgIleAr 62
Db 794 TCCGGCCGTGGGACGCGGCGTCTGTCGACGACGCGGAGCAATTCACGCTGTCTCCG 735
QY 62 gAspThr-----PhePheProArgPheHisGlyThrArgLeuLeuProThrG1 78
Db 734 GCGTACGGCGGCGCCAGTACCACATTCCGCGACGCTGTGTGA-----ATGACTGGCG 684
QY 78 uAlaGlnProGlyGluProHisProHisIleuValLeu----- 90
Db 683 AGAACAGCCCGGTGCCGGCGCGCGGCGCCGACCTCTGCTCGGTCCAGGCCAGCGATACGGC 624
QY 91 -----AspAspLeuLeuAlaGlyPheGlnAl 99

```


APPLICATION NUMBER: 2582 AGTCTCAGCGCCAACTGCAGGCGGTAGCCAGCACCTGCTGCTCTCGCGGTGCAGCGC 2523
FILING DATE: 186 sAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyValLeuSerGlnLeuArgGluLe 206
CLASSIFICATION: 2522 GGGGCTCCAGCGTGCAGCAGGCTGCCCTCGGGGCTCTGTCACCGA----- 2474
ATTORNEY/AGENT INFORMATION: 206 uLysAlaTrpPheGluGln---ThrLeuPheHisPheTyrSerAlaSerIleLeuLe 225
NAME: Zeller, Karen J. 2473 -CAGGTTGGCGCGCCAGCAGCTCTCTCTTGGTCACACAC-----CT 2430
REGISTRATION NUMBER: 37,071 225 uGlyTyrAspAlaAlaAlaAlaAlaGlyGlyAspGlyGly----- 239
REFERENCE/DOCKET NUMBER: PA-0002 US 2429 TGGGTTTTCGAGCGCCATCGCTTCATGTTGAGCGGCTACCGTGATGAGTACG 2370
TELECOMMUNICATION INFORMATION: 240 -GlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAs 259
TELEPHONE: (650) 855-0555 2369 CGGTCTCAGGCTGCAAGTTTGAT---GACCATCTCTCTG-----GGCATCGGCCA 2322
TELEFAX: (650) 845-4166 259 pHisAsnPheLeu-----GlyGlyLeuCysSerLeuIleLysPheValSe 274
INFORMATION FOR SEQ ID NO: 1132: 2321 GCATGACGTCTTGTATGCTGGCGCGCGCGGCTCGCGCTCTCCATGCGCAGTGTAGT 2262
SEQUENCE CHARACTERISTICS: 274 rAspIleValProGluThrProHisThrGlnProLeuGly 287
LENGTH: 4078 base pairs 2261 GGACTGTGTAGCGCGGATCTGGCGGTCTGCGCGCGGG 2222
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g1407624
US-09-016-434-1132

Alignment Scores:
Pred. No.: 0.104 Length: 4078
Score: 107.50 Matches: 81
Percent Similarity: 33.53% Conservative: 31
Best Local Similarity: 24.25% Mismatches: 112
Query Match: 7.07% Indels: 110
DB: 4 Gaps: 17

US-10-042-894A-8 (1-289) x US-09-016-434-1132 (1-4078)

QY 16 AlaSerAlaSerLysLeuGlyProLeu---IleAspGlySerGlyLeuPheTyrLysPro 34
DB 3077 AGTCTTCGCGGAGACTTGTCCCGCAGACGTCGCGTAGCGGA----- 3033
QY 35 LeuGlnAlaGlyAspArgGlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHis 54
DB 3032 ---CGGGGGGCTGAAGGGGCGCAGCGCCCGCGCGGTGTGGGCTCGCATTTGGAGGTAT 2976
QY 55 -----AlaAlaValProAla-ArgIleArgAspThrPhePheProArgPheHi 70
DB 2975 AGCGCGTGTGCGGCTTCAGCGCTCGAGCGTGTGAGCGTTCT----- 2936
QY 70 sGlyThrArgLeuLeuProThrGluAlaGlnPro----- 81
DB 2935 -----CGCGCGCGGCTCAGCGGCTCGCGGAGCTCAGTCTCTCGGG 2892
QY 82 -----GlyGluProHisProHisLeuValLeuAspAspLeuAlaG1 96
DB 2891 CAGGGCCAGGCGACCGGCTCCCGCAGCGCCACCGTGTATTGAGAT----- 2843
QY 96 yPheGlnAlaProCysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSe 116
DB 2842 -----GGCCCCGTGCGCTCGCGCGGCGGCGGTCAGCAGCGGA 2802
QY 116 rProGluProTyrIleAlaLysCysLeuAlaLysAspArgGlyThrSerValLeuLe 136
DB 2801 GAAGGAC-----GGTCCCGCGGAGGCGGTTCGCGCGCGCTCCA 2763
QY 136 uGlyPheArgValSerGlyValArgValVal----- 146
DB 2762 GAATCTCGGGGTGCCACGCGGCGGTCTCTCCGGGATGCTCAGGACCTCGGCTGCTCCT 2703
QY 147 -----GlyProGluGlyAlaValTrpArgThrGluArgProGluVally 161
DB 2702 CGCCAGGCGCGCGGCTCGGGCGCGCAGCGCGGAACATACGTGGCCCTTGTGCA 2643
QY 161 sAlaMetAspThrAlaGlyVal-----ArgAr 170
DB 2642 CGCTGATCCGTGTAGCGGTCTTCGAGGCGCGGAACCTCCAGGTGGCCAGGGGCGTGC 2583
QY 170 gValLeuArgArgTyrValSerSerValAlaAsp-----GluGlyMetAspCy 186
DB 170 -----GluGlyMetAspCy 186

```
Pred. No.: 0.24 Length: 6000
Score: 106.50 Matches: 80
Percent Similarity: 33.64% Conservative: 31
Best Local Similarity: 24.24% Mismatches: 117
Query Match: 7.00% Indels: 102
DB: 1 Gaps: 15

US-10-042-894A-8 (1-289) x US-08-348-006B-6 (1-6000)

QY 16 AlaSerAlaSerLysLeuGlyProLeu---IleAspGlySerGlyLeuPheTyrLysPro 34
Db 3077 AGTTCTTGGCGAGACTTGGTCCCGAGGACGTCGCGTAGCGGA----- 3033
QY 35 LeuGlnAlaGlyAspArgGlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHis 54
Db 3032 ---CGGGGGGGCTGAAGGGGCCAGGGCCCGGGCCCGCGGCTCGCACTTGGAGGTCT 2976
QY 55 -----AlaAlaValProAla-ArgIleArgAspThrPhePheProArgPheHi 70
Db 2975 AGCCCGTGTGGGCTTCAGGCCCTGCAGCGTGCAGCGTTCCT----- 2936
QY 70 sGlyThrArgLeuLeuProThrGluAlaGlnPro----- 81
Db 2935 -----CCGGCCCGGCTCAGCGGCTCAGCGTGCAGCGTTCCT----- 2936
QY 82 -----GlyGluProHisProHisLeuValLeuAspLeuLeuAlaGln 96
Db 2891 CAGGGCCCGAGGACCGCGCTCCCGCACGCGCACCGTATTGACGAT----- 2843
QY 96 yPheGlnAlaProCysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSe 116
Db 2842 -----GGCCCCCTTGGCTCGCGGCGGACGCGGTGGCAGCCAGCGGA 2802
QY 116 rProGluProTyrIleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeu 136
Db 2801 GAAGGAC-----GGTCCCGCGGAGGGGTTCGCGCGCGCTCCA 2763
QY 136 uGlyPheArgValSerGlyValArgValVal----- 146
Db 2762 GAATCTCGCGGTGGCCACGCGGGGTGTCTCCGCGGATGCTCAGGACCTCGGCTCGCTCCT 2703
QY 147 -----GlyProGluGlyAlaValTrpArgThrGluArgProGluVally 161
Db 2702 CGCCAGGCGCGCGGCTCGGGCCGCAAGCCGGAACATACGTGGCCCTTGTGCA 2643
QY 161 sAlaMetAspThrAlaGlyVal-----ArgAr 170
Db 2642 CGCTGTATGCGGTAGCGGTCTCCGAGGGCGGGAACCTCCAGGTGGCCAGGGCGGTGCG 2583
QY 170 gValLeuArgArgTyrValSerSerValAlaAsp-----GluGlyMetAspCy 186
Db 2582 AGTCCTCACGGCCAAACTCGAGCGGTAGCCACACCTGGTCTCGCGGTGCGCAGCGG 2523
QY 186 sAlaLeuAlaAlaValTyrGlyGlyLysGlyValLeuSerGlnLeuArgGluLe 206
Db 2522 GGGGCTCCAGGTGCGCAGCAGCGTGCCTCGGGGTCTCTGTCACCGA----- 2474
QY 206 uLysAlaTrpPheGluGluGln---ThrLeuPheHisPheTyrSerAlaSerIleLeu 225
Db 2473 -CAGGGTGGCGGCCCGACGACTGCTCCCTGGTGTCAACACAC-----CT 2430
QY 225 uGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGly----- 239
Db 2429 TGGGTATTGCTCGAGCGCCATCGCCCTTCATGTTGTAGGCGGTACCGTATGAGTACG 2370
QY 240 -GlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAsp---GlyValI 258
Db 2369 CGGTCTCAGGTGCAAGTTGTGATGACCATCTCTCGGCGATCGCGCAGCATGACGTCT 2310
QY 258 eAspHisAsnPheLeuGlyLysCysSerLeuIleLysPheValSerAspIleValPr 278
Db 2309 TGATCGCGGGCGCGCGGCGCTCGCGCCCTCCATGATCGGCACTAGTGGACCTGGTAGC 2250
```

```
QY 278 oGluThrProHisThrGlnProLeuGly 287
Db 2249 CGCGGATCGCGCGTGTCTCGCGCGCGG 2222

RESULT 7
US-08-800-825A-6/c
; Sequence 6, Application US/08900825A
; Patent No. 5866397
; GENERAL INFORMATION:
; APPLICANT: RODAN, GIDEON A.
; APPLICANT: SCHMIDT, AZRIEL
; APPLICANT: RUTLEDGE, SU JANE
; TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN
; TITLE OF INVENTION: TYROSINE PHOSPHATASE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSER: J. MARK HAND - MERCK & CO., INC.
; CITY: RAHWAY
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/800,825A
; FILING DATE: 14-FEB-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HAND, J. MARK
; REGISTRATION NUMBER: 36,545
; REFERENCE/DOCKET NUMBER: 18992DA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-3905
; TELEFAX: 732-594-4720
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6000 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-800-825A-6

Alignment Scores:
Pred. No.: 0.24 Length: 6000
Score: 106.50 Matches: 80
Percent Similarity: 33.64% Conservative: 31
Best Local Similarity: 24.24% Mismatches: 117
Query Match: 7.00% Indels: 102
DB: 2 Gaps: 15

US-10-042-894A-8 (1-289) x US-08-800-825A-6 (1-6000)

QY 16 AlaSerAlaSerLysLeuGlyProLeu---IleAspGlySerGlyLeuPheTyrLysPro 34
Db 3077 AGTTCTTGGCGAGACTTGGTCCCGAGGAACTGTCGCGTAGCGGA----- 3033
QY 35 LeuGlnAlaGlyAspArgGlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHis 54
Db 3032 ---CGGGGGGGCTGAAGGGGCCAGGGCCCGGGCCCGCGGCTCGCACTTGGAGGTCT 2976
QY 55 -----AlaAlaValProAla-ArgIleArgAspThrPhePheProArgPheHi 70
Db 2975 AGCCCGTGTGGGCTTCAGGCCCTGCAGCGTGCAGCGTTCCT----- 2936
QY 70 sGlyThrArgLeuLeuProThrGluAlaGlnPro----- 81
Db 2935 -----CCGGCCCGGCTCAGCGGCTCAGCGTGCAGCGTTCCT----- 2936
```

```
QY 82 -----GlyuProHisProHisLeuValLeuAspAspLeuAlaG1 96
Db 2891 CAGGCCCCAGGACCGCGCTCCCGCAGCGCCAGCGTGTATTGACGAT----- 2843
QY 96 yPheGlnAlaProCysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSe 116
Db 2842 -----GGCCCCGTTCGCTCGCGGGCAGCGGTGGCAGCCAGCGGA 2802
QY 116 rProGluProTyrIleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLe 136
Db 2801 GAAGGAC-----GGTCCCGCCGAGCGGTTCGCGCGCGCTCCA 2763
QY 136 uGlyPheArgValSerGlyValArgValVal----- 146
Db 2762 GAATCTCGCGGTGGCCAGCGCGGTCTCTCGGGATGCTCAGGACCTCGGCTGCTCCT 2703
QY 147 -----GlyProGluGlyAlaValTyrArgThrGluArgProGluValVal 161
Db 2702 CCGCCAGGCGCGCGGCTCCGGCCCGCAGCGGACACATACGTGGCCCTTGTGCA 2643
QY 161 sAlaMetAspThrAlaGlyVal-----ArgAr 170
Db 2642 CGCCTGATCGCGGTAGCGTCTTCGAGGGCGGGAATCCAGGTTGGCCAGGGCGTGC 2583
QY 170 gValLeuArgArgTyrValSerSerValAlaAsp-----GluGlyMetAspCy 186
Db 2582 AGTCCTCAGCGCCAACTCAGCGGTAGCCAGCAGCAGCTGTCCTCCGCGGTGGCCAGCG 2523
QY 186 sAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyValLeuSerGlnLeuArgGluLe 206
Db 2522 GGGGTCTCCAGCGTGCACAGAGCTCCCTCGGGGTCTGCTGCACCGA----- 2474
QY 206 uLysAlaTrpPheGluGluGln-----ThrLeuPheHisPheTyrSerAlaSerIleLeuLe 225
Db 2473 -CAGGTTTCGGCGGCCAGCACTGCTCCCTTGTGTACAAACCCAC-----CT 2430
QY 225 uGlyTyrAspAlaAlaValAlaGlyGlyAspGlyGly----- 239
Db 2429 TGGGTTGTCGAGCGCCATCGCTCCCTCATGTGTAGCGCGCTACCGTGTGGAGTAGC 2370
QY 240 -GlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAsp---GlyValI1 258
Db 2369 CGGTCTCAGCTGCAAGTTGTGTATGACCATCTCTCTGGGCATCGGCCAGCATGCTCCT 2310
QY 258 eAspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValPr 278
Db 2309 TGATGCGCGCGCGCGCGCTCGCGCGCTCCATGCGCAGTAGTGGACCTGTGTAGC 2250
QY 278 oGluThrProHisThrGlnProLeuGly 287
Db 2249 CCGGATCTGGCGGTGCTGTCGCGCGCGG 2222
```

RESULT 8

```
US-09-158-657-6/c
; Sequence 6, Application US/09158657
; Patent No. 6214564
; GENERAL INFORMATION:
; APPLICANT: RODAN, GIDEON A.
; APPLICANT: SCHMIDT, AZRIEL
; APPLICANT: RUTLEDGE, SU JANE
; TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN
; TITLE OF INVENTION: TYROSINE PHOSPHATASE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESS: J. MARK HAND - MERCK & CO., INC.
; STREET: 126 E. LINCOLN AVE., P.O. BOX 2000
; CITY: RAHWAY
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/158,657
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/800,825
; FILING DATE: 14-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: HAND, J. MARK
; REGISTRATION NUMBER: 36,545
; REFERENCE/DOCKET NUMBER: 18992DA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-3905
; TELEFAX: 732-594-4720
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6000 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-158-657-6

Alignment Scores:
Pred. No.: 0.24 Length: 6000
Score: 106.50 Matches: 80
Percent Similarity: 33.64% Conservativeness: 31
Best Local Similarity: 24.24% Mismatches: 117
Query Match: 7.00% Indels: 102
DB: 3 Gaps: 15
```

US-10-042-894A-8 (1-289) x US-09-158-657-6 (1-6000)

```
QY 16 AlaSerAlaSerLysLeuGlyProLeu---IleAspGlySerGlyLeuPheTyrLysPro 34
Db 3077 AGTTCTTGGCGAGACTTGTCCCGCAGGACGTCGCGTAGCGGA----- 3033
QY 35 LeuGlnAlaGlyAspArgGlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHis 54
Db 3032 ---CGGGGGGGGTGAAGGGGCCAGGGGCCCGCGCGCTGTGGCTCGCATTTGGAGTCA 2976
QY 55 -----AlaAlaValProAla-ArgIleArgAspThrPhePheProArgPheHi 70
Db 2975 AGCGCGTGTCCGGCTTCAGCCCTCGAGCTGAGCGTTCT----- 2936
QY 70 sGlyThrArgLeuLeuProThrGluAlaGlnPro----- 81
Db 2935 -----CGCGCGCGCGCTCAGCGCGCTCGCGCGCTCAGTCTCTCGGG 2892
QY 82 -----GlyGluProHisProHisLeuValLeuAspAspLeuAlaG1 96
Db 2891 CAGGCCCCAGGACCGCGCTCCCGCAGCGCCAGCGTGTATTGACGAT----- 2843
QY 96 yPheGlnAlaProCysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSe 116
Db 2842 -----GGCCCCGTTCGCTCGCGGGCAGCGGTGGCAGCCAGCGGA 2802
QY 116 rProGluProTyrIleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLe 136
Db 2801 GAAGGAC-----GGTCCCGCCGAGCGGTTCGCGCGCGCTCCA 2763
QY 136 uGlyPheArgValSerGlyValArgValVal----- 146
Db 2762 GAATCTCGCGGTGGCCAGCGCGGTCTCTCGGGATGCTCAGGACCTCGGCTGCTCCT 2703
QY 147 -----GlyProGluGlyAlaValTyrArgThrGluArgProGluValVal 161
Db 2702 CCGCCAGGCGCGCGCGGTCTCCGGCCCGCAGCGGACACATACGTGGCCCTTGTGCA 2643
QY 161 sAlaMetAspThrAlaGlyVal-----ArgAr 170
Db 2642 CGCCTGATCGCGGTAGCGTCTTCGAGGGCGGGAATCCAGGTTGGCCAGGGCGTGC 2583
```

```
Db 2642 CCGCTGATCCGCTGAGCGGTCTCCGAGAGGGGGAACTCCAGGGTGCACAGGGCGCTCG 2583
QY 170 gValLeuArgArgTyrValSerSerValAlaAsp-----GluglyMetAspCy 186
Db 2582 AGTCTCTACGGCCAACTGACAGCGGTAGCCAGCAGCAGCTGCTCCGCGGTGCAGCGC 2523
QY 186 salAlaLeuAlaAlaValTyrGlyGlyslgGlyValLeuSerGlnLeuArgGluLe 206
Db 2522 GGGGTCCAGCGTCCAGCGAGCGTCCCTCGGGGTCTGCTGCACCGA----- 2474
QY 206 ulysAlaTrpPheGluGluGln---ThrLeuPheHisPheTyrSerAlaSerIleLeuLe 225
Db 2473 -CAGGGTGGCGCGCCACACTGCTCCCTTGGTCAACACAC-----CT 2430
QY 225 uGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGly----- 239
Db 2429 TGGGTTTGTGCGAGCGCCATCGCCTTCATGCTGTAGCGGCTACCGTGTAGGAGTACG 2370
QY 240 -GlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAsp---GlyValIl 258
Db 2369 CCGTCTCAGGTCGCAAGTTTGTGATGACCATCTCCTGGGCATCGGCCAGCATGAGCTCT 2310
QY 258 easPheHisAsnPheLeuGlyGlyLeuCysserLeuIleLysPheValSerAspIleValPr 278
Db 2309 TGAATCGCGCGCGCGCGCGCTCGCGCCCTCCATCGCGCAGTGTAGTGGACCTGCTAGC 2250
QY 278 oGluThrProHisThrGlnProLeuGly 287
Db 2249 CCGGATCTGGCGCTGCTCGCGCGCGG 2222

RESULT 9
PCT-US94-10166-6/c
; Sequence 6, Application PC/TUS9410166
; GENERAL INFORMATION:
; APPLICANT: RODAN, GIDEON A
; APPLICANT: RODAN, GIDEON A
; APPLICANT: SCHMIDT, AZRIEL
; APPLICANT: RUTLEDGE, SU JANE
; TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN
; TITLE OF INVENTION: CDNA ENCODING A NOVEL HUMAN PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JOHN W. WALLEN III
; STREET: P.O. BOX 2000, 126 E. LINCOLN AVE.
; CITY: RAYWAY
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/10166
; FILING DATE: 09-SEPT-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/122,032
; FILING DATE: 14-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: WALLEN, JOHN W III
; REGISTRATION NUMBER: 35403
; REFERENCE/DOCKET NUMBER: 18992
; TELEPHONE: 908-594-3905
; TELEFAX: 908-594-4720
; TELEX: 138825
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6000 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
i MOLECULE TYPE: cDNA
PCT-US94-10166-6
Alignment Scores:
Pred. NO.: 0.24 Length: 6000
Score: 106.50 Matches: 80
Percent Similarity: 33.64% Conservative: 31
Best Local Similarity: 24.24% Mismatches: 117
Query Match: 7.00% Indels: 102
DB: 5 Gaps: 15

US-10-042-894A-8 (1-289) x PCT-US94-10166-6 (1-6000)
QY 16 AlaSerAlaSerLysLeuGlyProLeu---IleAspGlySerGlyLeuPheTyrLysPro 34
Db 3077 AGTTCTGGGGGAGACTTGTCCCGCAGGAACGTCGGTAGCGGA----- 3033
QY 35 LeuGlnAlaGlyAspArgGlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHis 54
Db 3032 ---CGGGGGGCTGAAGGGGCCAGGGCCCGCGCGCTGTGGGCTCGCACTTGGAGGTCT 2976
QY 55 -----AlaAlaValProAla-ArgIleArgAspThrPhePheProArgPheHi 70
Db 2975 AGGCGGTCTCGGGCTTCAGGCCCTGCAGCGCTGAGCGTTCT----- 2936
QY 70 sGlyThrArgLeuLeuProThrGluAlaGlnPro----- 81
Db 2935 -----CGCGCCCGGCTCAGCGCGCTCGCGGAGCTCAGTCTCTCGGG 2892
QY 82 -----GlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGl 96
Db 2891 CAGGGCCCGAGGACCGGGCTCCCGCAGCGCCAGCGTGTATTGACGAT----- 2843
QY 96 yPheGlnAlaProCysValAlaAspIleLysIleGlyAlaIleThrTrpProProSerSe 116
Db 2842 -----GGCCCCGTTGCGCTCGCGGGGCGACGGGTGGCAGCCAGCGGA 2802
QY 116 rProGluProTyrIleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLe 136
Db 2801 GAAGGAC-----GGTCCCGCGGAGCGGTTCGCGCGCGCTCCA 2763
QY 136 uGlyPheArgValSerGlyValArgVal----- 146
Db 2762 GAATCTGGGGTGGCGACGCGGGCTGTCTCCGGGATGCTCAGGACCTCGGCTGCTCT 2703
QY 147 -----GlyProGluGlyAlaValTrpArgThrGluArgProGluValVally 161
Db 2702 CGCCCGAGCGCGCGGGCTCCGGGGCGGCGAGCGGACACATACGTGGGCCCTTGTGCA 2643
QY 161 sAlaMetAspThrAlaGlyVal-----ArgAr 170
Db 2642 CGCCTGATGCCGTGTAGCGGTCTCCGAGGGCGGGAACCTCCAGGGTGGCGAGGGCGCTCG 2583
QY 170 gValLeuArgArgTyrValSerSerValAlaAsp-----GluGlyMetAspCy 186
Db 2582 AGTCTCTACGGCCAACTGACAGCGGTAGCCAGCAGCAGCTGCTCCGCGGTGCAGCGCG 2523
QY 186 salAlaLeuAlaAlaValTyrGlyGlyLysGlyValLeuSerGlnLeuArgGluLe 206
Db 2522 GGGGTCCAGCGTCCAGCGAGCGTCCCTCGGGGTCTGCTGCACCGA----- 2474
QY 206 ulysAlaTrpPheGluGluGln---ThrLeuPheHisPheTyrSerAlaSerIleLeuLe 225
Db 2473 -CAGGGTGGCGCGCCACACTGCTCCCTTGGTCAACACAC-----CT 2430
QY 225 uGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGly----- 239
Db 2429 TGGGTTTGTGCGAGCGCCATCGCCTTCATGCTGTAGCGGCTACCGTGTAGGAGTACG 2370
QY 240 -GlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAsp---GlyValIl 258
Db 2369 CCGTCTCAGGTCGCAAGTTTGTGATGACCATCTCCTGGGCATCGGCCAGCATGAGCTCT 2310
```

```
Qy 258 eAspHisAsnPheLeuGlyLeuCysSerLeuLeuLeysPheValSerAspIleValPr 278
Db 2309 TGATGCGCGCGCGCGCGCGCGCTCGCGCCCTCATCGGCACGATAGTGGACCTGGTAGC 2250
Qy 278 oGluThrProHisThrGlnProLeuGly 287
Db 2249 CGCGGATCTGCGCGTGTGCGCGCGCG 2222

RESULT 10
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37RV
US-09-103-840A-1

Alignment Scores:
Pred. No.: 5,48+03 Length: 4411529
Score: 106.00 Matches: 78
Percent Similarity: 33.88% Conservative: 25
Best Local Similarity: 25.66% Mismatches: 106
Query Match: 6.97% Indels: 97
DB: 3 Gaps: 16

US-10-042-894A-8 (1-289) x US-09-103-840A-1 (1-4411529)
Qy 11 ValAlaGlyHisArg-----AlaSerAlaSerLysLeuGlyProLeuLeaSpGly 27
Db 1951504 CTCGCGCGCCACCGTGAACGATAGCCCGCCGCTGTTCGGGACCGCTTCGATGAT 1951563
Qy 28 SerGlyLeuPheTyLys-----ProLeu-----GlnAlaGlyAspArgGlyGlu 42
Db 1951564 GCGGTGCTGTCTGAAGTCGCGGTAGCGCTGTGGCCGACGACCGAGCGGAGATCCCGGCGAG 1951623
Qy 43 HisGluValAlaPheTyGluAlaPheSerAlaHisAlaAlaValProAlaArgIleArg 62
Db 1951624 GCCCGCTCGATTCTGTGAACCGCA-GCGGCGCGAGAGAGTCCCGCACGATTAGCA 1951682
Qy 63 AspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGly 82
Db 1951683 GGTGCC-----AGCGGTCCCCGACGAGCTCCGCGCGCGCGCTACGCGGC 1951727
Qy 83 GluPro-----HisProHisLeuValLeuAsp 91
Db 1951728 AGTACTGGCGGTACGCGCTGACGACACGAGGTAGTCGCGCATCCCTGGCTGGTGTTCG 1951787
Qy 92 AspLeuLeuAlaGlyPheGlnAlaProCysValAlaAspIleLysIleGlyAla----- 109
Db 1951788 GAATTC-----GAACTTCCCGACCCCTGTGGGAGCGGTACGCTTGTCTGTGAGGT 1951841
Qy 110 -----IleThrTyrProProSer-----SerProGlu 118
Db 1951842 GAGAGCGGATGACCGGACGCTGACCAAGACGCTGGGTTCCCTCGACGATTTCAGGGAA 1951901
Qy 119 ProTyrIleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPhe 138
Db 1951902 CGCTTTGTGTCCTCCCGGTGATCCGGAATACCCAGGG----- 1951937
```

```
Qy 139 ArgValSerGlyValArgValValGlyProGluGly---AlaValTyrArgThrGluArg 157
Db 1951938 -----TGCGGCGCATCTGGAACGGGSCAGGTGCGCCGCGAACCGG 1951976
Qy 158 Pro--GluValLysAlaMetAspThrAlaGlyValArgValLeuArgTyrVal 176
Db 1951977 CTTTGATCGCCACGTCGCAC-GACGCGTGGCATGTCCGAACGGTGTCTCGCGCGCGGTG 1952035
Qy 177 SerSerValAlaAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGly--- 195
Db 1952036 -----GACGCGGAGTGTGACCGCGGTACGTGGCGCGCGG 1952071
Qy 196 -----LysGlyGlyValLeuSerGlnLeuArgGluLeu 206
Db 1952072 CACAACGTGGCGGCACCGCGCTGTCGACGCGCGGTGGTGTGATCGACCTCTCGCGCGATG 1952131
Qy 207 LysAlaTyrPheGluGluGlnThrLeuPheHisPheTyrSerAlaSerIleLeuLeuGly 226
Db 1952132 CGGCGC-----
Qy 227 TyrAspAlaAlaAlaValAlaAlaGlyAspGlyGlyValThrValLysLeu--- 245
Db 1952144 CTGATCCAGCGACTGGCGCGGTACGGTGCAGGGTGTGTCACCGCTCGCGGATTGGAC 1952203
Qy 246 -----ValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHisAsnPhe 262
Db 1952204 CACGCCACGGTCCCGTTCGCGCGGTGCGCCCGCC-----GGGATCGTCACCAACCGCGT 1952260
Qy 263 LeuGlyGlyLeu 266
Db 1952261 GTCGCGCGGCTG 1952272

RESULT 11
US-08-311-731A-132/c
; Sequence 132, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 132:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36412 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
```

```
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: MYCOBACTERIUM LEPRAE
US-08-311-731A-132

Alignment Scores:
Pred. No.: 4.65 Length: 36412
Score: 105.50 Matches: 74
Percent Similarity: 35.33% Conservative: 32
Best Local Similarity: 24.67% Mismatches: 119
Query Match: 6.94% Indels: 75
DB: 4 Gaps: 14

US-10-042-894A-8 (1-289) x US-08-311-731A-132 (1-36412)
QY 6 ProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyPro----- 23
DB 16731 CCGGATGAATAGCCATTGCACCTAGAAACATCAAGCAAGCGCCGAAAGGCCCTGCTGCA 16672
QY 24 -----LeuIleAspGlySerGly---LeuPhe 31
DB 16671 TGAACACCTTGACGGTGGTGGCCCGACGCGGTGTTGATATTCCCGCCGAGGTGGC 16612
QY 32 TyrLysProLeuGlnAlaGlyAspArgGlyGluHisGluValAlaPheTyrGluAlaPhe 51
DB 16611 TATGATAGGCTCCCGCTACCGACGTTGAGTCGCTAGACATGGTTTCGCACTGCT--- 16555
QY 52 SerAlaHisAlaValProAlaArgIleArgAspThrPhePheProArgPheHisGly 71
DB 16554 ---TCGACAGCGCGCTGTTGGAGCGCTACCTG----- 16525
QY 72 ThrArgLeuLeuProThrGluAlaGlnProGlyGluProHisProHisLeuValLeuAsp 91
DB 16524 -----GAGCGGTTTCTCATACCGTG----- 16504
QY 92 AspLeuLeuAlaGlyPheGlnAlaPro-----CysVal 102
DB 16503 -----GCCGTGATGCAGACACAGAGCGGTTGCACCGGCTGCTTATGAGTGGTG 16453
QY 103 AlaAspIleLysIleGlyAlaIleThrTrpProPro-----SerSerProGluProTyr 120
DB 16452 GAAGACCTCGCCGCTGATTGGTGGTCTAGCCGCAAGTCGGGTTCGCTCGGAGCTGCAC 16393
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
DB 16392 ATCGACGAGGGTGTGTTGATGAGGCTCTGCAAGCGTGTGGGGGTTT---GCT 16336
QY 141 SerGlyValArgValValGlyProGluGly---AlaValTrpArgThrGluArgProGlu 159
DB 16335 GACGGTGAAAGGGCTTGTGCGCGGAGGCAATGCCATC-----ACAGTGGCTGCTTG 16282
QY 160 ValLysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSerVal 179
DB 16281 GTACAGCGATGCGGACGCGGCAATGTCGGGAGATCCCGAGTAGCATCCGGTTC 16222
QY 180 AlaAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGly--- 198
DB 16221 CGGACAAAGGATC-----GTCGGATTGACATCGCTGCTGCTGAGGCTGGGCAC 16171
QY 199 -----ValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGlu 211
DB 16170 CTTCCGACTCGGACCTCGACGCTTTTGTAGTACATGCTAGTAAATATGCGCGCTTC--- 16114
QY 212 GluGlnThrLeuPheHisPheTyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaA 231
DB 16113 -----ACTATTTCATCGGGTGGAGCGGCTGCGGCTGCCATCTATCCAGGCGCATC 16063
QY 232 ValAlaAlaGlyGlyAspGlyGlyGlyValThrValLysLeuValAspPheAlaHisVal 251
DB 16062 GCGTTCTCGGAGCTGACCGGCTGGGCCACGGGTGGGATCGTCGACGACATCGAGTT 16003
QY 252 AlaGluGlyAspGlyValIleAspHisAsnPheLeuLeuGlyGlyLeuCysSerLeuIleLys 271
DB -----ATTACGATGCAGACCCAGAGCTGTACGAGAG 1265
```

```
Db 16002 GATCCTGGTGGTCATC-----CGGTTGGTCCCTGGCTCGATCCTGGCG 15955

RESULT 12
US-08-637-899-2
; Sequence 2, Application US/08637899
; Patent No. 5908772
; GENERAL INFORMATION:
; APPLICANT: Mitta, Masanori
; APPLICANT: Sano, Mutsumi
; APPLICANT: Kato, Ikunoshin
; TITLE OF INVENTION: Gene Encoding Lacto-N-Biosidase
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch and Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/637,899
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiner, Marc S
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 1422-252P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1917 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Streptomyces sp.
; STRAIN: 142
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1917
; IDENTIFICATION METHOD: E
US-08-637-899-2

Alignment Scores:
Pred. No.: 0.0826 Length: 1917
Score: 104.00 Matches: 66
Percent Similarity: 34.73% Conservative: 25
Best Local Similarity: 25.19% Mismatches: 97
Query Match: 6.84% Indels: 74
DB: 2 Gaps: 13

US-10-042-894A-8 (1-289) x US-08-637-899-2 (1-1917)
QY 4 LeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyPro 23
DB 1149 CTCATCCAGCAGCC---GTCCTCGCTGCTCGCCGAGGCGCGCC----- 1190
QY 24 LeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis 43
DB 1191 -----GGTCATGAATCCGCGCTACCTCTACCTGGTGGCGCGCGG 1232
QY 44 GluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgIleArgAsp 63
DB 1233 -----ATTACGATGCAGACCCAGAGCTGTACGAGAG 1265
```

```
QY 64 ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83
Db 1266 -----CGACTGACGCCGCTTGGCTTCAGGGGCGAGCGCTGACCCAGGGGCG 1313
QY 84 ProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaProCysValAla 103
Db 1314 GCGCAACCTCACCGCGCGAGATCAGCTGCTGGCGCGA----- 1352
QY 104 AspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyrlleAlaLys 123
Db 1353 -----CAG 1355
QY 124 CysLeuAlaLysAsp-----ArgGlyThrThrSerValLeuLeuGlyPheArgValSer 141
Db 1356 TCGCGCGCGAGAGCGGAGACAGGTTCGAGACGAGGTCTTCATGCCGCTCGCTTCGT 1415
QY 142 GlyValArgValValGly-----ProGluGlyAlaValTrpArgThrGlu 156
Db 1416 GCGCGAGCGACCTGGGGCGGCCGAGCGCGAGCCGACGTACGCGGCTTCGAGGCGCT 1475
QY 157 ArgProGluValLysAlaMetAspThrAlaGlyValArgArg-----ValLeuArg 173
Db 1476 CCGCGGAGATCGGTACGCGCGCGGCTGGGAGAACACCGACCGCGCGCTCGCCGA 1535
QY 174 ArgTyrlleSerSerValAlaAspGluGlyMetAspCysAlaLeuAlaAlaValTrp 193
Db 1536 CGGTACGTACCG-----GCTGACACGCGCGCGA-----GGCGCTGGCGCCAC 1580
QY 194 GlyGlyLysGlyGlyValLeuSerGlnLeuArgGlu-----LeuLysAlaTrpPheGlu 211
Db 1581 GCGCGACGCGCGCTGTCTCTGTCTGTAAGAACAGCGCGGCTCTGGGCGCTGACGCGAC 1640
QY 212 GluGlnThrLeuPheHis-----PheTyrlleSerAlaSerIleLeuLeuGly----- 226
Db 1641 CCGCGACGCGGTACTACCGCTGCTCCACGAGAGCGGTCTGCTGCGCGCGGTGCG 1700
QY 227 ---TyrlleAspAlaAlaAlaValAlaAlaGlyGlyAspGlyGlyValThrValLysLeu 245
Db 1701 CGGCAAGAGTACTGGGGCGCGCGCTGAGGTTGGGGGCGGAGCTGTACTCGCAACTG 1760
QY 246 Valasp 247
Db 1761 CTCGAC 1766

RESULT 13
US-09-252-991A-11331
; Sequence 11331, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11331
; LENGTH: 1896
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11331

Alignment Scores:
Pred. No.: 0.0925 Length: 1896
Score: 103.50 Matches: 105
Percent Similarity: 34.54% Conservative: 29
Best Local Similarity: 27.06% Mismatches: 112
Query Match: 6.80% Indels: 145
```

```
DB: 4 Gaps: 23
US-10-042-894A-8 (1-289) x US-09-252-991A-11331 (1-1896)
QY 16 AlaSerAlaSerLysLysLeuGlyProLeuIleAspGlySerGlyLeuPheTyrllys-Pro-- 34
Db 716 TCCGCGCGCCAGCAAACTG-----GCGAGCGGTTGGCGACGAAAGTCCGGG 760
QY 35 -LeuGlnAlaGly---AspArgGlyGluHisGluVal----- 45
Db 761 TCGCGGACAGCCCTCGATTCGCGGTAGCACGCGCGACGCTCCAGCACGTAATGGGCG 820
QY 46 -AlaPheTyrlleGluAlaPheSerAlaHisAlaValPro----- 58
Db 821 CGCTGTACCGCTCGGCTTCTGCCACGCGAGCTTTCAGCATCTCGAGGCTCTCGCGC 880
QY 59 -----AlaArgIleArg-AspThrPhe-----PheProArgP 69
Db 881 AGGCGCAGCCGCGCGCGCGCGGTGCGTGTATCAGTTCGCGCGCGTTCGCCACG 940
QY 69 heHisGlyThrArgLeuLeu-----ProThrGluAlaGlnProGlyGluProHisP 86
Db 941 TTCACCTCAGCGCGCGGTTCGATCACCGGTAGTTCGCGCAGGTTCGCGCGC----- 993
QY 86 roHisLeuValLeuAspAspLeuLeuAlaGlyPheGln-----AlaProC 101
Db 994 -----GCTCCAGATGCGCGCGCGCGCGGACCCCGGTGAGCAGCTTGTCCGCT 1042
QY 101 ysValAlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluPProTyrl 121
Db 1043 GTCTCGCTTCGTAGAA-----ACGATCCGCGCGGTAGCGCGGAGGCGTTC 1090
QY 121 leAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal- 140
Db 1091 AGGTACGTGCA-CTCGGTTGCGGAGCAGGTTCATTCGTGGGATTCGCGCTT 1149
QY 141 -----SerGlyValArgValValGlyPro-----GluGlyAlaValTrpA 154
Db 1150 GTCGAGGAGATCCGCGCGGTTTCGGGTCTTCGCGATCGCTTCAGCGCA---TGCG 1206
QY 154 rgThrGluArgProGluValLysAlaMetAspThrAlaGlyValArgArgValLeuArgA 174
Db 1207 GCGCGGTACAGGTAG-ATCCGTCGACGAGTACCCATCCAC-CGCCAGCGGATCGCG 1264
QY 174 rgTyrl-----ValSerSerValAla----- 180
Db 1265 GGCACACATGTCGCGCAGGGGCTTTCGTTCGTAGCGGCTCTCAATTCGACCGCGC 1324
QY 181 -----AspGluGlyMetAspC 186
Db 1325 GCGCGCAGCGCGGATCGCGCGCGCGCGCGGTTCGACGAGCAGCTGGGCTCGACC 1384
QY 186 ysAlaLeuAlaAlaValTyrlGlyGlyLysGlyVal----- 199
Db 1385 TTCCCAT---GCGCGGTACATGTCGCGTAGGCGGCTTCGCGCGCGCTCGCGGGG 1441
QY 200 -----LeuSerGlnLeuArgGluLeuLysAlaTrpPheGluGlnThrLeuP 216
Db 1442 TCGAGGACGATAGTTCGGCTGCGCACCGGCTTCGCGAGAGGAAGATCCACCGCG 1501
QY 216 heHisPheTyrlSerAlaSerIleLeuLeuGlyTyrl----- 227
Db 1502 CCGAG-----TCCGAGCGGTAGGCTTCGCGTACCGCGCGGCGAGCGCGATGCCGCC 1555
QY 228 -----AspAlaAlaAlaValAlaAlaGly----- 235
Db 1556 GCAGCGTCAGGCGTTTCCGCCATTGGCCAGGTTTCCAGCGCGCGCGCGTTCGCGGCC 1615
QY 236 -----GlyAspGlyGly-----GlyValThrValLysLeuValA 247
Db 1616 GGATGTGGCGTGGCGACGCGCGGCTTCGCGGTGCGTGGGCGTTCGCGGAAAGTGTG 1675
QY 247 spPheAlaHisValAlaGlu-----GlyAspGlyV 257
```

```
Db 1676 GCATGCGCGCGAGGGTGTAGCAGGAGCAGCGCGCAGCGGACAGGCTACGAGCGGC 1735
      |||||
QY 257 alileaspHisasnPheLeuGlyGlyLeuGlyCysSerLeuLeuLeuLysPheValSerAspIleV 277
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 1736 AGGTTGAGGAAGGACATGCGCGGCTCACTGTTTCAGCGCGCTGTCGCGGAGCGGG- 1788
      |||||
QY 277 alProGluThrProHisThr 283
      |||||
Db 1789 --CCGGAACGCTCAGACT 1806
      |||||

RESULT 14
US-09-252-991A-12565/c
; Sequence 12565, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12565
; LENGTH: 960
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12565

Alignment Scores:
Pred. No.: 0.0724 Length: 960
Score: 100.50 Matches: 68
Percent Similarity: 37.60% Conservative: 23
Best Local Similarity: 28.10% Mismatches: 83
Query Match: 6.61% Indels: 70
DB: 4 Gaps: 9

US-10-042-894A-8 (1-289) x US-09-252-991A-12565 (1-960)

QY 80 GlnProGlyGluProHisProHisLeuValLeu----- 90
Db 759 CAGCCCGGTCCGCGCGCGCGACCTCGTCTGCTCCAGCGCGGATACAGCGCGG 700
QY 91 -----AspAspLeuLeuAlaGlyPheGlnAlaPro 100
Db 699 CTTGAGCGCGCGAGATACTGTCATGCGGCGAGCATTTCTGCGCGGTATAGCAGCCCT 640
QY 101 CysValAlaAsp-----11e-Ly 106
Db 639 CTTGAGAGCTGACCCCGCGCGACCGCTGCGAGTTGATCATCTGTGGGATGACAGCTCGCG 580
QY 106 sileGlyAlaIleThrTriProProSerProGluProTyIleAlaLysCysLeuAl 126
Db 579 GGTAGCGCGGAACACCTTGCGG-----ATGCTTCG-CGCGACCTGGCGGAGCGCAAT 527
QY 126 aLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyValArgValVa 146
Db 526 CGTCGAGCGCGCTCGCGGCTGTGTCGC-GGAGCGGTGGCAAGCACCGCGCTCGCGCGCT 468
QY 146 lGlyProGluGlyAlaValThrArgThrGluArgPro--GluValLysAla-MetAspT 165
Db 467 TCGCGCGGAACCCCATAGCTACGCGCGGCGCTGCCCGAGGCGCGAGCGAGCGTCCGCG 408
QY 165 hrAlaGlyValArgValLeuArgValLeuArgValVal----- 176
Db 407 TGGCGGCGCATCTCGCGCTCTCGCGCGGGGTGTCTATCCCGAGGCGCGCAGTGTGCTCG 348
QY 177 -----SerSerValAlaAspGluGlyMetAspCysAlaLeuAlaAlaValT 193
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
```

```
Db 347 CTGGCGTCCGCGCAAGCCTATGCGGAGCCAGGC-----TGCCTCTCGTCCGCAAGCCTG 294
QY 193 yRGlyGlyLysGlyGlyValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGluGluG 213
      |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 293 GCCTTGGAGATACCGCGTATTTCCTCAGGTCCGCCACTGGGCGCTC-----GAGC 243
QY 213 lThrLeuPheHisPheTyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaA----- 231
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 242 AGCTCGCGCGCATCGCCAGCAGCGCGCTCGCCCTCGGGCAGGATCGGAGCTCGAT 183
QY 232 -----ValAlaAlaGlyGlyAspGlyGlyValThrValLysLeuV 246
Db 182 AGCAGCGCGCGCTTTCATGTTGAGCGCGCGCGAGCTGAGGTCTCGTCGTTGAG----- 127
QY 246 alaspPheAlaHisValAlaGluGlyAspGlyValIleAspHisasnPheLeuGlyGly 266
Db 126 -----GTAGCGAGGTTCAGGTGCGCTGCGCGCTCGGAGCACTTGGCGCG 81
QY 266 eu 266
      ::
Db 80 TC 79

RESULT 15
US-09-144-914-3/c
; Sequence 3, Application US/09144914
; Patent No. 6309855
; GENERAL INFORMATION:
; APPLICANT: Duprat, Fabrice
; APPLICANT: Lesage, Florian
; APPLICANT: Fink, Michel
; APPLICANT: Lazdunski, Michel
; TITLE OF INVENTION: FAMILY OF MAMMALIAN POTASSIUM CHANNELS, THEIR CLONING
; TITLE OF INVENTION: AND THEIR USE, ESPECIALLY FOR THE SCREENING OF DRUGS
; FILE REFERENCE: 989.6705CIP
; CURRENT APPLICATION NUMBER: US/09/144,914
; CURRENT FILING DATE: 1998-09-01
; EARLIER APPLICATION NUMBER: 08/749,816
; EARLIER FILING DATE: 1996-11-15
; EARLIER APPLICATION NUMBER: 60/095,234
; EARLIER FILING DATE: 1998-08-04
; EARLIER APPLICATION NUMBER: FR 96/01565
; EARLIER FILING DATE: 1996-02-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2514
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (126)..(1307)
; FEATURE:
; OTHER INFORMATION: TASK
US-09-144-914-3

Alignment Scores:
Pred. No.: 0.307 Length: 2514
Score: 100.50 Matches: 71
Percent Similarity: 35.12% Conservative: 34
Best Local Similarity: 23.75% Mismatches: 108
Query Match: 6.61% Indels: 87
DB: 4 Gaps: 14

US-10-042-894A-8 (1-289) x US-09-144-914-3 (1-2514)

QY 10 GlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyProLeuLeuAspGlySerGly 29
Db 1259 CAGACCGGTGACACCGAGCT-----CATGGC---GGA 1230
QY 30 LeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHisGluValAlaPheTyrGlu 49
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 1229 GCGTGCGCCGCCCTGCTGCACAGCAGCGCTGCGAGGCGGTGTCGCTGTAGCGCGCCGCC 1170
```



```
QY 50 AlaPheSerAlaHisAlaAlaValProAlaArgIleArgAspThrPhePheProArg--- 68
Db 1169 TCCCGCGACGAGTGGCTCTGTCTCCACGCACGTGCGAGAGTCCCGCGGAT 1110
QY 69 PheHisGly-----ThrArgLeuLeuProThrGluAlaGlnPro 81
Db 1109 GATCATGGGGATGGAGTACTGACAGCTTCTCGGGGCTCTTGACCAAGCAGCAGCACAT 1050
QY 82 GlyGluProHisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaProCys 101
Db 1049 GGACTGGAAGTCAGCACCTCCGCGTA---GACGTTGGGAAGCCGCCCGCGCTGC 993
QY 102 -ValAlaAspIleLysIleGlyAlaIle---ThrTrpProProSerSerProGluProTyr 120
Db 992 CGCCGTGATGAGCGGTGTCGTAGTGTGCGGTGCGCCCTCCCGCGCCCGCCGC 933
QY 120 rife-----AlaLysCysLeuAlaLysAspArgGlyThrSerVa 134
Db 932 CTGCCCCGTGGCGGTGAGCAGCGCGGTGTCGGCG---TCGCGCTTCTCGTCTCGGC 876
QY 134 IleLeuGlyPheArgValSerGlyValArgValGlyProGluGlyAlaValTrpAr 154
Db 875 GTTCATGTGTCATGAAGCGCAGCACCGAGGTGAG-----GAAGCGCGCGATGACCG 823
QY 154 gThrGluArgProGluValLysAlaMetAspThrAlaGlyValArgArg----- 170
Db 822 TGAGCCCGCTAGGATAGTAGAAGCTGAAGCCACCGTACTGCGGCTGCGTCTCAGGG 763
QY 171 -----ValLeuArgArgTyrValSerSerValAlaAspGluGlyMetAspCysAlaLe 188
Db 762 CCTGTCTCTTGCAGCGCCACGTAGTCCCGAAGCCGAT----- 723
QY 188 uAlaAlaValTyrGlyGlyGlyGlyValLeuSerGlnLeuArgGluLeuLysAl 208
Db 722 -----GTTGGTGAGGGT----- 711
QY 208 aTrpPheGluGluGlnThrLeuPheHisPheTyrSerAlaSerIleLeuLeuGlyTyrAs 228
Db 710 -----GATGAAGCAGTAGTAGTAGGCTGGA 685
QY 228 pAlaAlaAlaValAlaAlaGlyAspGlyGlyGlyValThrValLysLeuValAspPh 248
Db 684 AGAGGTCCAGTGTCTGTAGTCGGAGAAGCGCGCGCCGATGCACACGCTGCTGATGC 625
QY 248 eAlaHisValAlaGluGlyAspGlyValIleAspHisAsnPheLeuGlyGlyLeuCysSe 268
Db 624 ACGAGAGAAGCCGAT--GAGCACCATGTTGGCCATGGACACGTCGCGC----- 579
QY 268 rLeuIleLysPheValSerAspIleValProGluThrProHisThrGlnProLeu 286
Db 578 -----GCGCGCATGCCCGCCCTT 558
```

Search completed: March 27, 2004, 09:02:16
Job time : 2142 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 27, 2004, 04:36:55 ; Search time 477 Seconds
(without alignments)
2573.856 Million cell updates/sec

Title: US-10-042-894A-8

Perfect score: 1521

Sequence: 1 MSDLHPPEHQVAGHRASASK.....IKFVSIVPEPHTOPLGPS 289

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10042894/runat_24032004_145046_16937/app_query.fasta_1.455
-DB=N Geneseq 29Jan04 -QMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTENT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10042894_QCGN_1_1_470 @runat_24032004_145046_16937 -NCPU=6 -ICPU=3
-NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 29Jan04:*

1: geneseq1980s:*
2: geneseq1990s:*
3: geneseq2000s:*
4: geneseq2001as:*
5: geneseq2001bs:*
6: geneseq2002s:*
7: geneseq2003as:*
8: geneseq2003bs:*
9: geneseq2003cs:*
10: geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1521	100.0	1344	6 AAD43514	Aad43514 Maize ino
2	1485	97.6	922	6 AAD43513	Aad43513 Maize ino
3	1435	94.3	923	6 AAD43512	Aad43512 Maize ino
4	1406	92.5	1169	6 AAD43511	Aad43511 Maize ino
5	1307	85.9	3416	6 AAD43522	Aad43522 Maize ino
6	1233	81.1	899	6 AAD43518	Aad43518 Maize ino
7	808.5	53.2	643	6 AAD43519	Aad43519 Maize ino
8	742.5	48.8	1020	6 AAD43517	Aad43517 P. argent

9	722.5	47.5	1195	6	AAD43516	Aad43516 Eucalyptu
10	714.5	47.0	1130	3	AAC48750	Aac48750 Arabidops
11	700.5	46.1	1130	3	AAC40313	Aac40313 Arabidops
12	674.5	44.3	1104	3	AAC33685	Aac33685 Arabidops
13	674.5	44.3	1168	3	AAC39023	Aac39023 Arabidops
14	674.5	44.3	1243	3	AAC38692	Aac38692 Arabidops
15	644.5	42.4	1105	6	AAD43515	Aad43515 Soybean i
16	464	30.5	519	6	AAD43520	Aad43520 Maize ino
17	329	21.6	353	6	AAD43521	Aad43521 Maize ino
18	306.5	20.2	484	9	ADE82058	Ade82058 Arabidops
19	288	18.9	2930	4	ABL23306	Ab123306 Drosophil
20	273.5	18.0	464	6	ABL93575	Ab193575 Arabidops
21	265	17.4	876	4	ABL23307	Ab123307 Drosophil
22	201.5	13.2	2212	7	AAD53176	Aad53176 Human kin
23	192	12.6	2391	9	ADC06742	Adc06742 Human ino
24	192	12.6	2890	6	AAD38858	Aad38858 Human kin
25	191	12.6	2608	7	ADA53453	Ada53453 Human cod
26	189	12.4	1340	4	AAI60607	Aai60607 Human pol
27	189	12.4	1737	5	AAH64793	Aah64793 Human sec
28	189	12.4	1737	9	ADC06746	Adc06746 Human ino
29	189	12.4	1739	4	AAH16813	Aah16813 Human CDN
30	189	12.4	1748	4	AAI58821	Aai58821 Human pol
31	189	12.4	1748	8	ADB48802	Adb48802 Novel hum
32	189	12.4	1757	5	AAH64800	Aah64800 Human sec
33	189	12.4	1990	3	AAH64898	Aah64898 Human sec
34	183	12.0	1918	3	AAC77195	Aac77195 Human ORF
35	182.5	12.0	1222	9	ADC06748	Adc06748 Murine in
36	182	12.0	1863	7	ABX70876	Abx70876 Novel hum
37	178.5	11.7	1893	7	ABX70875	Abx70875 Novel hum
38	168	11.0	1723	9	ADC06744	Adc06744 Human ino
39	168	11.0	4461	6	ABK83730	Abk83730 Human CDN
40	156	10.3	6223	4	ABL05232	Ab105232 Drosophil
41	151	9.9	3197	4	ABL05244	Ab105244 Drosophil
42	145.5	9.6	1782	6	ABL65412	Ab165412 Lung canc
43	145.5	9.6	1782	6	ABL64777	Ab164777 Lung canc
44	145	9.5	1386	6	ABL56444	Ab156444 Nucleotid
45	143	9.4	1241	4	AAI87016	Aai87016 Human pol

ALIGNMENTS

RESULT 1
AAD43514
ID AAD43514 standard; DNA; 1344 BP.
XX
AC AAD43514;
XX
DT 14-NOV-2002 (first entry)
XX
DE Maize inositol polyphosphate kinase (IPPK) DNA #4.
XX
KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
KW nutritional value; animal feed; transgenic; gene; ds.
XX
OS Glycine max.
XX
FH Key Location/Qualifiers
FT CDS 52..921
FT /*tag= a
FT /product= "Maize IPPK protein #4"
XX
FN WO200259324-A2.
XX
PD 01-AUG-2002.
XX
PF 09-JAN-2002; 2002WO-US003120.
XX
PR 12-JAN-2001; 2001US-0261465P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX

Alignment Scores:

Pred. No.: 1,07e-127 Length: 922
Score: 1485.00 Matches: 283
Percent Similarity: 98.27% Conservative: 1
Best Local Similarity: 97.92% Mismatches: 5
Query Match: 97.63% Indels: 0
DB: 6 Gaps: 0

US-10-042-894A-8 (1-289) x AAD43513 (1-922)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerIlys 20
DB 53 ATGCCCGACCTCCACCCCGGAGCACCAGTCGCGGCTCCCGCGCTCCGCGCAG 112
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
DB 113 CGGGCCCGCTCATCGAGGCTCCGGCTCTTCTACAAGCGCTCCAGCGCGGACCGT 172
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
DB 173 GGGAGACAGAGTGGCTTCTATGAGCGTTCCTCCGCCACGCGCGTCCCGCGCGC 232
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
DB 233 ATCCGAGACACCTTCTCCCGCGGTCCACGCGACGACTCTCCCGACCGGCGCG 292
QY 81 ProGlyGluProHisProHisValLeuAspLeuAlaGlyPheGlnAlaPro 100
DB 293 CCGGGAGCGCATCCGACCTCTCTCGAGACCTCTCCGGGATTTGAGCGGCC 352
QY 101 CysValAlaAspIleLysIleGlyAlaIleThrProProSerSerProGluProTyr 120
DB 353 TCGGTCCGAGACATCAAGATCGGCCCATCAGTGGCCACCGAGTTCGCGGAGCCCTAC 412
QY 121 IleAlaLysCysLeuAlaIysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
DB 413 ATCGCAAGTGGCTCGCATGACCGCGGACCGAGCGTCTCTCGGATTCGCGCTC 472
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
DB 473 TCCGGCTCCGAGTCTCGGCCCGGAGCGCGCTGTGCGGAGCGGCGCGAGGTG 532
QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSerValAla 180
DB 533 AAGGCCATGGACACCGCGCGCTCCGCGCGTCTCCGCGCTACCTGTCTCCGTTGCC 592
QY 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyLysGlyGlyValLeu 200
DB 593 GACGAGGGGATGGACTGTGCGCTCGCGCGGCTACGAGGAGGAAAGGTGGAGTCTTG 652
QY 201 SerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSer 220
DB 653 TCACAGCTGGCGAGCTCAAGCGTGGTTCGAGGAGCAGACTCTGTTCACCTTACTCG 712
QY 221 AlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGly 240
DB 713 GGTCTGATTTCTTGGGCTATGATGCTGCTGAGTCCGAGTCGAGCGGCGAGGTGGG 772
QY 241 ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHis 260
DB 773 GTGACGGTGAACCTGGTGGACTTTCCTCATGTGGCGGAGGTGGGTGATTACCCAC 832
QY 261 AsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluThr 280
DB 833 AACTTCTCGGCGGCTCTGCTCGCTGATCAAGTTCGATTCGATTCGATTCGAGACT 892
QY 281 ProHisThrGlnProLeuGlyProSer 289
DB 893 CCTCAGACCGCCTTTGGGTCTCTTCT 919

RESULT 3

AAD43512

ID AAD43512 standard; DNA; 923 BP.

XX AAD43512;
AC
XX 14-NOV-2002 (first entry)
DT
XX Maize inositol polyphosphate kinase (IPPK) DNA #2.
DE
XX Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
KW nutritional value; animal feed; transgenic; gene; ds.
XX
XX Zea mays.
OS
XX
XX Key Location/Qualifiers
FT CDS 53..736
FT /*tag= a
FT /product= "Maize IPPK protein #2"
XX
XX WO200259324-A2.
FN
XX 01-AUG-2002.
PD
XX 09-JAN-2002; 2002WO-US003120.
PF
XX 12-JAN-2001; 2001US-0261465P.
PR
XX (PION-) PIONEER HI-BRED INT INC.
PA
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
PI WPI; 2002-636540/68.
XX P-PSDB; AAE26194.
DR
XX
PT New inositol polyphosphate kinase polynucleotides and polypeptides,
PT useful in modulating phytic acid biosynthesis by decreasing phytate or
PT increasing non-phytate phosphorous to improve the nutritional value of
PT animal feed.
XX
PS Claim 1; Page 60-61; 86pp; English.
XX
CC The invention relates to novel inositol polyphosphate kinase (IPPK)
CC polypeptides and polynucleotides. Sequences of the invention are useful
CC in modulating the phytic acid biosynthesis by decreasing phytate and/or
CC increasing non-phytate phosphorous to improve the nutritional value of
CC animal feed, or to reduce the environmental impact of animal waste.
CC Polynucleotides of the invention are used to produce transgenic plants with an
CC altered phenotype. IPPK proteins are used to screen compounds that
CC modulate their activity and raising anti-idiotypic antibodies. The
CC present sequence is maize IPPK DNA
XX
SQ Sequence 923 BP; 138 A; 326 C; 286 G; 173 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.41e-123 Length: 923
Score: 1435.00 Matches: 277
Percent Similarity: 96.21% Conservative: 2
Best Local Similarity: 95.52% Mismatches: 10
Query Match: 94.35% Indels: 1
DB: 6 Gaps: 0

US-10-042-894A-8 (1-289) x AAD43512 (1-923)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerIlys 20
DB 53 ATGCCCGACCTCCACCCCGGAGCACCAGTCGCGGCTCCCGCGCTCCGCGCAG 112
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
DB 113 CGGGCCCGCTCATCGAGGCTCCGGCTCTTCTACAAGCGCTCCAGCGCGGACCGT 172
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
DB 173 GGGAGACAGAGTGGCTTCTATGAGCGGCTTCTCCGCCACGCGCGTCCCGCGCGC 232

Db 562 AAGGCTATGACACCGTCCGGCGTCCGGCGTCTCCGGCGCTAGCTGTCTCCGGCTTGC 621
Qy 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyGlyGlyValle 200
Db 622 CGACGAGGGATGGACTCGCGCTCCGGCGCGGTGTACGAGGAGAAAGGTGGAGTCTT 681
Qy 200 uSerGlnLeuArgGluLeuLysAlaTyrPheGluGlnThrLeuPheHisPheTyrSe 220
Db 682 GTACACAGCTGCCGAGCTCAAGCGATGGTGGAGAGGAGACTCTGTCCACTTCTATC 741
Qy 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyAspGlyGly 240
Db 742 GGCGTCGATTCTCTGGGCTATGATGCTGCTGCACTCCGAGCGGAGGTGGGGTGG 801
Qy 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValleAspHi 260
Db 802 GGTAACAGTGAAGCTGGTGGACTTGTCCCATGTGCGCGAGGTGATGGGGTGAATGACA 861
Qy 260 sAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280
Db 862 CAACCTCTCTGGCGAGCTCTGCTAGCTGATCAAGTTCGTTCTGACATTGTTCCAGAGAC 921
Qy 280 rProHisThrGlnProLeuGlyProSer 289
Db 922 TCCTTAGACGACGCTTTGGGTCTTCT 949
RESULT 5
AAD43522
ID AAD43522 standard; DNA; 3416 BP.
XX
AC AAD43522;
XX
DT 14-NOV-2002 (first entry)
XX
DE Maize inositol polyphosphate kinase (IPPK) DNA #9.
XX
KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
KW nutritional value; animal feed; transgenic; gene; ds.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT 72..407
FT CDS
FT /*tag= a
FT /product= "Maize IPPK protein #6"
XX
PN WO200259324-A2.
XX
PD 01-AUG-2002.
XX
PF 09-JAN-2002; 2002WO-US003120.
XX
PR 12-JAN-2001; 2001US-0261465P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX
DR WPI; 2002-636540/68.
DR P-PSDB; AAE26201.
XX
PT New inositol polyphosphate kinase polynucleotides and polypeptides,
PT useful in modulating phytic acid biosynthesis by decreasing phytate or
PT increasing non-phytate phosphorous to improve the nutritional value of
PT animal feed.
XX
PS Claim 1; Page 75-77; 86pp; English.
XX
CC The invention relates to novel inositol polyphosphate kinase (IPPK)
CC polypeptides and polynucleotides. Sequences of the invention are useful
CC in modulating the phytic acid biosynthesis by decreasing phytate and/or
CC increasing non-phytate phosphorous to improve the nutritional value of
CC animal feed, or to reduce the environmental impact of animal waste.

CC Polynucleotides of the invention are to produce transgenic plants with an
CC altered phenotype. IPPK proteins are used to screen compounds that
CC modulate their activity and raising anti-idiotypic antibodies. The
CC present sequence is maize IPPK DNA
XX

SQ Sequence 3416 BP; 893 A; 830 C; 856 G; 837 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,45e-110 Length: 3416
Score: 1307.00 Matches: 260
Percent Similarity: 90.34% Conservative: 2
Best Local Similarity: 89.66% Mismatches: 9
Query Match: 85.93% Indels: 20
DB: 6 Gaps: 1

US-10-042-894A-8 (1-289) x AAD43522 (1-3416)

Qy 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20
Db 72 ATGCCCGACCTCCACCCGCGGAGCACCAAGTCGCGGTCCCGCGCTCCGCCAGCAAG 131
Qy 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
Db 132 CTGGGCCCACTCATCGACGACTCTGGCCTCTTCTAAGCCGCTCCAGCCCGGACCGT 191
Qy 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArg 60
Db 192 GGGAGACAGAGGTGGCTTCTATGAGGGGTCTCCGCCACGCGCTCCGCCCGCC 251
Qy 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 252 ATCCGAGACACCTTCTCCCGGTTCACGCGACGACTCTCCGCCACCGCGCGCAG 311
Qy 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro 100
Db 312 CCGGGAGCGCGCATCCGACCTCTGCTCGACGACCTCTCTCGGGGTTCGAGCGGCC 371
Qy 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyr 120
Db 372 TCGCTCGCAGACATCAAGATCGTGCATCAGTG----- 406
Qy 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 407 -----ACCACGAGCGTTCGCTCGGATTCCGCGTC 436
Qy 141 SerGlyValArgValValGlyProGluGlyAlaValTyrArgThrGluArgProGluVal 160
Db 437 TCCGGCGTCCGAGTCTCGGGCCCGAGGCGCGGTGTGCGGACGAGCGCGCGAGGTG 496
Qy 161 LysAlaMetAspThrAlaGlyValArgValLeuArgValLeuArgTyrValSerSerValAl 180
Db 497 AAGGCTATGGACATTGTGGCGTCCGCCGCGTCTCCGCCGCTAGCTGTCTATCCGTTGC 556
Qy 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyGlyGlyValle 200
Db 557 CGACGAGGGATGGACTCGCGCTCCGGCGCGGTGTACGAGGAGAAAGGTGGAGTCTT 616
Qy 200 uSerGlnLeuArgGluLeuLysAlaTyrPheGluGlnThrLeuPheHisPheTyrSe 220
Db 617 GTCACAGCTCGCGAGCTCAAGCGGTGGTTCAGGGGCGAGACTCTGTTTCACTTCTACTC 676
Qy 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyAspGlyGly 240
Db 677 GGCGTCGATTCTTCTGGGCTATGATGCTGCTCAGTCCGACGAGCGGAGGTGGGGTGG 736
Qy 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValleAspHi 260
Db 737 GGTAACAGTGAAGCTGGTGGACTTTGCCCATGTGCGCGAGGTGATGGGGTGAATGACA 796
Qy 260 sAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280
Db 797 CAACCTCTCTGGCGGCGCTCTGCTAGCTGATCAAGTTTGTTCGACATTGTTCCAGAGAC 856

```

QY 280 rProHisThrGlnProLeuGlyProSer 289
DB 857 TCCTCAGAGCGACCTTTGGGTCCTTCT 884

RESULT 6
ID AAD43518
AC AAD43518;
XX 14-NOV-2002 (first entry)
DE Maize inositol polyphosphate kinase (IPPK) DNA #5.
DE Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
KW nutritional value; animal feed; transgenic; gene; ds.
XX Zea mays.
XX
XX Location/Qualifiers
FH Key 89..424
FT CDS /*tag= a
FT /product= "Maize IPPK protein #4"
XX
XX WO200259324-A2.
XX 01-AUG-2002.
XX
XX 09-JAN-2002; 2002WO-US003120.
XX
XX 12-JAN-2001; 2001US-0261465P.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX WPI; 2002-636540/68.
XX P-PSDB; AAE26200.
XX
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
XX useful in modulating phytic acid biosynthesis by decreasing phytate or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed.
XX Claim 1; Page 73-74; 86pp; English.
XX
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
XX polypeptides and polynucleotides. Sequences of the invention are useful
XX in modulating the phytic acid biosynthesis by decreasing phytate and/ or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed, or to reduce the environmental impact of animal waste.
XX Polynucleotides of the invention are to produce transgenic plants with an
XX altered phenotype. IPPK proteins are used to screen compounds that
XX modulate their activity and raising anti-idiotypic antibodies. The
XX present sequence is maize IPPK DNA
XX
XX Sequence 899 BP; 140 A; 312 C; 276 G; 171 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.87e-104 Length: 899
Score: 1233.00 Matches: 248
Percent Similarity: 88.42% Conservative: 4
Best Local Similarity: 87.02% Mismatches: 14
Query Match: 81.07% Indels: 20
DB: 6 Gaps: 1

US-10-042-894A-8 (1-289) x AAD43518 (1-899)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerLys 20
DB 89 ATGCCGACCTCCACCCGCGGAGACCAAGTCGCCGGTCACCGCGCTCCGCGACCAAG 148
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40

```

```

DB 149 CTGGCCCCACTCATCGAGCGCTCTGGCTCTTCTATAAGCGCTCCAGGCCGCGACCGT 208
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
DB 209 GGGGAGCAGAGGTGCGCTTCTATAGCGGTCTCCGCCACCGCGCGCTCCCGGCCCGC 268
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
DB 269 ATCCGAGACACTTCTTCCCGCGGTTCACGGCAGCGACTCTCTCCCGCAGGCGCAG 328
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro 100
DB 329 CCGGGGAGCGCATCTCCGTACCTCGCTCGACGACCTCTCGCGGGGTTTGAGGCGCC 388
QY 101 CysValAlaAspIleLeuGlyAlaIleThrTyrProProSerSerProGluProTyr 120
DB 389 TCGTCCGAGACATCAAGATCGTGCCATCAGTCACCATG-AGCGAT----- 435
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
DB 436 -----CTGCTCGGATTCCACGTC 453
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTyrArgThrArgProGluVal 160
DB 454 TCCGGCGTCCGAGTCTGCGGCCCGCGCGCGGTGTGCGGAGCGGCGCGCTGAGGTG 513
QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSer-ValAl 180
DB 514 AAGGCTATGGACATGTGCGGCGTCCCGCGGTGTCCGGCGTGCATGTCTCCGTTGC 573
QY 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyVal 200
DB 574 CGCGGAGGGGATGGACTCGCGCTCGCGCGCGGTGTACGGAGGAGGAGGTGGAGTCTT 633
QY 200 uSerGlnLeuArgGluLeuLysAlaTyrPheGluGlnThrLeuPheHisPheTyrSe 220
DB 634 GTCACAGCTGCGCGAGCTCAAGGCGTGTTCGAGGGGCGAGCTCTGTTCACCTTCTACTC 693
QY 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGly 240
DB 694 GCGGTCGATTTCTTGGGCTATGATGCTGTGCACTCCGACGAGCGGAGGTGGGGGTGG 753
QY 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260
DB 754 GGTAACAGTGAAGCTGGTGGACCTTGGCCCATGTGGCGAGGGGTGATGGGTGATTGACCA 813
QY 260 sAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280
DB 814 CAACTTCCTGGCGGGCTCTGCTAGTATCAAGTTTGTTCGACATTGTTCAGAGAGAC 873
QY 280 rProHisThrGln 284
DB 874 TCCTTAGCGCAG 886

RESULT 7
AAD43519
ID AAD43519 standard; DNA; 643 BP.
XX
XX AAD43519;
XX
XX 14-NOV-2002 (first entry)
XX
XX Maize inositol polyphosphate kinase (IPPK) DNA #6.
XX
XX Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
XX nutritional value; animal feed; transgenic; ds.
XX
XX Zea mays.
XX
XX WO200259324-A2.
XX
XX 01-AUG-2002.

```

```
XX 09-JAN-2002; 2002WO-US003120.
PF 12-JAN-2001; 2001US-0261465P.
PR (PION-) PIONEER HI-BRED INT INC.
PA Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
PI WPI; 2002-636540/68.
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
XX useful in modulating phytic acid biosynthesis by decreasing phytate or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed.
PS Claim 1; Page 74; 86pp; English.
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
XX polypeptides and polynucleotides. Sequences of the invention are useful
XX in modulating the phytic acid biosynthesis by decreasing phytate and/ or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed, or to reduce the environmental impact of animal waste.
XX Polynucleotides of the invention are to produce transgenic plants with an
XX altered phenotype. IPPK proteins are used to screen compounds that
XX modulate their activity and raising anti-idiotypic antibodies. The
XX present sequence is maize IPPK DNA
XX
SQ Sequence 643 BP; 92 A; 261 C; 178 G; 103 T; 0 U; 9 Other;

Alignment Scores:
Pred. No.: 1.85e-65 Length: 643
Score: 808.50 Matches: 162
Percent Similarity: 93.14% Conservative: 1
Best Local Similarity: 92.57% Mismatches: 10
Query Match: 53.16% Indels: 3
DB: 6 Gaps: 0

US-10-042-894A-8 (1-289) x AAD43519 (1-643)
QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerIys 20
Db 117 ATGCCCGACCTCCACCCCGGAGACCAAGTCGCGCTACCCGCGCTCCGCCAGCAG 176
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
Db 177 CTGGCCCGCTCATCGAGGCTCCGGCTCTTCTACAGCGCTCCAGCGCGGACCGT 236
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
Db 237 GGGAGACGAGGTGGCTTCTATGAGCGTTCCTCGGCCACCGCGNGCTCCCGCGCGC 296
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 297 ATCCGAGACACCTCTTCCCGCGGTTCACGCGACCGGACTCTCCACCGAGCGCGAG 356
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100
Db 357 CCGCGGAGCGCATCCGACCTCTCTCGACGACTCTCTCGCGGGTTTGAGCGGCC 416
QY 101 CysValAlaAspIleIysIleGlyAlaIleThrTrpProProSerSerProGluProTyr 120
Db 417 TCGGTTCGAGACATCAAGATCGCGCCCATCATCGTGGCCACCGAGTTCGCGGAGCCCTAC 476
QY 121 IleAlaIleCysLeuAlaIleAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 477 ATCGNCAAGTACTTNGCCAGGACCGCGGACCAAGAGGTTCTCTCGGATTCGCGGTTC 536
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTyrArgThrGluArg-ProGluVal 160
Db 537 TTGC--GTCCGAGTGTGCGGCCCGGAGGCGCGGTGTGGCGGACGAGCGGCCCGGGGT 594
QY 160 IlysAlaMetAspThr-AlaGlyValArgArgValLeuArg 173
```

```
Db 595 GAANGCTATGACACCCGTCGGNGNCCGGCGNGTGCTTCGG 635
RESULT 8
RAD43517
ID AAD43517 standard; DNA; 1020 BP.
XX AAD43517;
XX 14-NOV-2002 (first entry)
XX P. argentatum inositol polyphosphate kinase (IPPK) DNA #1.
XX Enzyme; inositol polyphosphate kinase; IPPK; phytic acid; transgenic;
XX nutritional value; animal feed; gene; ds.
XX Parthenium argentatum.
XX
FH Key Location/Qualifiers
FT CDS 21..908
FT /*tag= a
FT /product= "P. argentatum IPPK protein #1"
XX
PN WO200259324-A2.
XX 01-AUG-2002.
XX 09-JAN-2002; 2002WO-US003120.
XX 12-JAN-2001; 2001US-0261465P.
XX (PION-) PIONEER HI-BRED INT INC.
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX WPI; 2002-636540/68.
XX P-PSDB; AAE26199.
XX
New inositol polyphosphate kinase polynucleotides and polypeptides,
XX useful in modulating phytic acid biosynthesis by decreasing phytate or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed.
XX Claim 1; Page 71-72; 86pp; English.
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
XX polypeptides and polynucleotides. Sequences of the invention are useful
XX in modulating the phytic acid biosynthesis by decreasing phytate and/ or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed, or to reduce the environmental impact of animal waste.
XX Polynucleotides of the invention are to produce transgenic plants with an
XX altered phenotype. IPPK proteins are used to screen compounds that
XX modulate their activity and raising anti-idiotypic antibodies. The
XX present sequence is P. argentatum IPPK DNA
XX
SQ Sequence 1020 BP; 297 A; 195 C; 219 G; 309 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4.04e-59 Length: 1020
Score: 742.50 Matches: 149
Percent Similarity: 67.93% Conservative: 45
Best Local Similarity: 52.10% Mismatches: 79
Query Match: 48.82% Indels: 13
DB: 6 Gaps: 5

US-10-042-894A-8 (1-289) x AAD43517 (1-1020)
QY 4 LeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerIysLeuGlyPro 23
Db 24 CTCAGGGCCCCAGATCATCATGCTTGTGACATGAAGCTGGGCTCGGAGAGCTTGGCCCA 83
QY 24 LeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis 43
```



```
Db 84 CTCATTGATGATTACAGCGCGGTTTTACAAACACATCGCAGGCGTATACCGTGGGTGACAA 143
QY 44 GluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgIleArgAsp 63
Db 144 GAAGTAGCCTTTTAAATCATTTCTTCTTCAATATATATTTCCAGAACACATACGC--- 200
QY 64 ThrPhePheProArgPheHisGlyThrArgLeuLeuLeuProThrGluAlaGlnProGlyGlu 83
Db 201 AAATTTCTTCTATATATATGACCAAAATCATG-----AAGGCATCCACTGGCTCT 254
QY 84 ProHisProHisLeuValLeuAspIleLeuAlaGlyPheGlnAlaProCysValAla 103
Db 255 GACCATCTCTCATCATGCTGTGCAAGATCTTACATCAGCTCATGTCAACCCATCTGTAAAG 314
QY 104 AspIleLeuValAlaIleThrTrpProProSerSerProGluProTyrIleAlaLys 123
Db 315 GACATCAAAATCGGTCCAGACATCGGCGCCAGAGCTTCGAGCGGTACATTGCMAAA 374
QY 124 CysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyVal 143
Db 375 TGCTTAAAAAGGATAGGAAAGACACAGTATTTCCATTGGGATTCAGAGATCTCCGGGTG 434
QY 144 ArgValValGlyProGluGlyAla---ValTyrArgThrGluArgProGluValLysAla 162
Db 435 CAAGCTATATCATGATATGGTCAGCGGTTTTATAGCCCTCATAGAAATTACATGCGTAAA 494
QY 163 MetAspThrAlaGlyValArgValLeuArgArgTyrValSerSerValAlaAspGlu 182
Db 495 ACCGGCCAGCTCATGTAGACTCTTCTTAGGAATTTGTTCTTCTAACCCGCTCGCA 554
QY 183 -----GlyMetAspCysAlaLeuAlaAlaValTyrGly 194
Db 555 GAGATGGAATGCGCACAGCGCTAGCGCGGATTTCTTTAGCATCTTTTGTATGTGT 614
QY 195 GlyGlyGlyGlyValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThr 214
Db 615 GGCCCTAATGGATATAGTCAACTGATGATGATGATGATGATGATGATGATGATGATGAT 674
QY 215 LeuPheHisPheTyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAla 234
Db 675 ATTTACCACCTCTATGCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 734
QY 235 GlyGlyAspGlyGlyValThrValLysLeuValAspPheAlaHisValAlaGluGly 254
Db 735 GGTGCT---CGGTCAACACGACAGAGTCAAACTTATTTGATTTTCTCATGTATACAGATGT 791
QY 255 AspGlyValIleAspHisnProLeuGlyGlyLeuCysSerIleLeuLysPheValSer 274
Db 792 AATGGTGTTATGATCACAAATTTCTTGGGTGGGCTCTGTCTTTGATTAAGTTTCATTTCT 851
QY 275 AspIleValProGluThr 280
Db 852 GACATACTTTCGGAGACA 869
RESULT 9
ID AAD43516
XX standard; DNA; 1195 BP.
AC AAD43516;
XX
XX 14-NOV-2002 (first entry)
XX Eucalyptus grandis inositol polyphosphate kinase (IPPK) DNA.
DE DE
XX Enzyme; inositol polyphosphate kinase; IPPK; phytic acid; transgenic;
KW nutritional value; animal feed; gene; ds.
XX
XX Eucalyptus grandis.
OS
XX Key Location/Qualifiers
FH 116..1048
FT CDS /*tag= a
FT /product= "E. grandis IPPK protein"
```

```
XX WO200259324-A2.
FN
XX 01-AUG-2002.
XX
XX 09-JAN-2002; 2002WO-US003120.
XX
XX 12-JAN-2001; 2001US-0261465P.
XX
XX (PION-) PIONEER HI-BRED INT INC.
PA
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
PI WPI; 2002-636540/68.
XX P-PSDB; AAE26198.
DR
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
PT useful in modulating phytic acid biosynthesis by decreasing phytate or
PT increasing non-phytate phosphorous to improve the nutritional value of
PT animal feed.
XX
XX Claim 1; Page 68-70; 86pp; English.
XX
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
CC polypeptides and polynucleotides. Sequences of the invention are useful
CC in modulating the phytic acid biosynthesis by decreasing phytate and/or
CC increasing non-phytate phosphorous to improve the nutritional value of
CC animal feed, or to reduce the environmental impact of animal waste.
CC Polynucleotides of the invention are used to produce transgenic plants with an
CC altered phenotype. IPPK proteins are used to screen compounds that
CC modulate their activity and raising anti-idiotypic antibodies. The
CC present sequence is Eucalyptus grandis IPPK DNA
XX
XX Sequence 1195 BP; 342 A; 240 C; 294 G; 319 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 3 456-57 Length: 1195
Score: 722.50 Matches: 149
Percent Similarity: 66.31% Conservative: 38
Best Local Similarity: 52.84% Mismatches: 84
Query Match: 47.11 Indels: 6
DB: 6 Gaps: 6
US-10-042-894A-8 (1-289) x AAD43516 (1-1195)
QY 4 LeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyPro 23
Db 119 CTCAAGGTCCTCGGATCATCAAGTCGCGGTCAACGCGGAGACGGGGGAAAGCTGGGCGCA 178
QY 24 LeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis 43
Db 179 CTGGTGGATGATTCGGGCGGCTTCTATAGCTCTCCAGAGCGATCATCGGAGACAGC 238
QY 44 GluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgIleArgAsp 63
Db 239 GAAGTGGCTTTTACGAGTCATTCTTATCCATCCGAGATCCCGAGTCCACATTCGC--- 295
QY 64 ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83
Db 296 AAATTTCTTCTCTCGGTTTCCGGAACATAAGACTATT-----GAGCGCTCTGTATGATCG 349
QY 84 ---ProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaProCysVal 102
Db 350 GGTCTTCAACCTCACCTGTTCTGGAGATCTGCTCGGTCGACGACCAACCATCTCTC 409
QY 103 AlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyrIleAla 122
Db 410 ATCGACATCAAGACTGGATCCAGAACATGTTATCGGAGGCGCTCTCAGAGTACATCCAA 469
QY 123 LysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGly 142
Db 470 AAGTGCTTAGAAGAAAGATCGAAATAGCAACAGCGTTTCTATTGGGTTTCTAGGATTTCTGG 529
```

```
QY 143 ValArgValVal---GlyProGluGlyAlaValTIPArgThrGluArgProGluVallys 161
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 530 CTAAGGGTATACAAATAGCAGCTGATTTGGCAACCTGAGAGAGAGGTTGTTAT 589
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 162 AlaMetAspThrAlaGlyValArgArgValLeuArgTyrValSerValAlaAsp 181
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 590 AGCTTTAATGCGGAGCTGAGTGGCTCTGAGGAGTTGTTCTTCCAACTTGCT 649
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 182 GluGly-----MetAspCysAlaLeuAlaAlaValTyrGlyGly 197
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 650 CTGGGTCCAAATGTGATCCGATTTGTTGTATGATCAAAAGTTTACTGTCCCGGGT 709
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 198 GlyValLeuSerGlnLeuArgGluLeuLysAlaTTPheGluGluGlnThrLeuPheHis 217
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 710 GGAATTTGGCACATTTGCTCAGCTGAGGAGTGGTTGAGGTTGAGGTTGAGGTTGAG 769
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 218 PheTyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAsp 237
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 770 TTCTATTTCTTGTCTACTCATTTATATATGACAGGGAGTCTGTTGACGGC----- 823
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 238 GlyGlyGlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyVal 257
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 824 TTGTGACACCCGAAAGTTAACTGGTGGACTTTGCATGTGATGATGCCACGGCGG 883
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 258 IleAspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleVal 277
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 884 ATCGATCAACAATCTTGTGGTGGCTCTGTTCTGTAATCAAGTTTATAGTCATGCT 943
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 278 ProGlu 279
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 944 GATGAA 949

RESULT 10
AAC48750
ID AAC48750 standard; DNA; 1130 BP.
XX
AC AAC48750;
DT 18-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 58629.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134258P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144336P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145087P.
```

```

PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145218P.
PR 26-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 23-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.

PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161040P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 1,76e-56 Length: 1130
Score: 714.50 Matches: 150
Percent Similarity: 66.55% Conservative: 39
Best Local Similarity: 52.82% Mismatches: 82
Query Match: 46.98% Indels: 13
DB: 3 Gaps: 8

US-10-042-894A-8 (1-289) x AAC48750 (1-1130)
QY 4 LeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerIysLeuGlyPro 23
Db 94 CTCAAGGTCCTCGAACACCAAGTTGCTGTCACATTGCTAGTGATGGGAAGCTCGGTCCA 153
QY 24 LeuIleAspGlySerGlyLeuPheTyrIysProLeuGlnAlaGlyAspArgGlyGluHis 43
Db 154 CTCGATGATGACCAAGCGCGTTCTTCAAGCACCTTCAGGGAGATTCTCGTGGCAACAC 213
QY 44 GluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArgIleArgAsp 63
Db 214 GAGGCTAAGTTCTATGAGTCTTTCATCGAACATGAAGGTTCCAGATCACATCCAT--- 270
QY 64 ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83
Db 271 AGATACTTCCCGGTGTATCACGGCCTCAGCTAGTT-----GAAGCATCTGATGATCT 324
QY 84 ProHis---ProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaProCysVal 102
Db 325 GCGAAGCTTCTCATCTGTTCTTGTATGATGTTGTTGAGGTACGCAAAACCGTCGGTA 384
QY 103 AlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyrIleAla 122
Db 385 ATGGATGTTAAGATTGGATCTAGACATGGTACCGGATGTATCAGAAGAACTTCAAG 444
QY 123 LysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGly 142
Db 445 AAATGTTATTAAAGAGATAGACAGACCACCGTTTCTGGGGTTTCAAGGTTTCAGGT 504
QY 143 ValArgValValGly---ProGluGlyAlaValTrpArgThrGluArgProGluValLys 161
Db 505 TTTAAGATTTTGTATCACCACCAAGATCAAGTTTGTGAGAGCTGAGAAGAAGCTTGTCT 564
QY 162 AlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSer-----SerVal 179
Db 565 GGGTATAATGCAGATGGTGTCTAGATTGGCTCTGAGAGAGTTTGTGTATCGCAACTCTCC 624
QY 180 AlaAspGluGlyMet-----AspCysAlaLeuAlaAlaValTyrGlyGlyLysGly 197
Db 625 GCTGACTCTAACTTGACCAACACTGTGCTTTTGTGATCAGAGGTTTATCGGGTTTAAAC 684
QY 198 GlyValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHis 217
Db 685 GGGATCTTAGCGCAGTTGTTGGAGCTTAAAGATTGGTTCGAAACCAACCGCTTACCAT 744
QY 218 PheTyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaVal---AlaAlaGlyGly 236
Db 745 TTCAATTCTCGATTCTGTATGATTTTACGAGAAATGAATCAATCTTGTGCAAGGAGA 804
```

Qy 237 AspGlyGly-----GlyValThrVallysLeuValAspPheAlaHisValAlaGlu 253
 Db 805 GATGATGACCGGCACCGGCACCAAGTGAAGCTGGTGGATTTCGCGCATGTTCTTGAT 864
 Qy 254 GlyAspGlyValIleAspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheVal 273
 Db 865 GGAAACGGGTGTCATCGACCAATAATTCTTGGTGGACTCTGCTCTTTTCATAAAGTTTCATC 924
 Qy 274 SerAspIleVal 277
 Db 925 AAAGATATTCTT 936
 RESULT 11
 AAC40313
 ID AAC40313 standard; DNA; 1130 BP.
 AC AAC40313;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 27858.
 XX
 KW Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway; metabolic pathway;
 KW promoter; termination sequence; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-00301439.
 XX
 XX 25-FEB-1999; 99US-0121825P.
 PR 05-MAR-1999; 99US-0123180P.
 PR 09-MAR-1999; 99US-0123548P.
 PR 23-MAR-1999; 99US-0125788P.
 PR 25-MAR-1999; 99US-0128264P.
 PR 29-MAR-1999; 99US-0128785P.
 PR 01-APR-1999; 99US-0127462P.
 PR 06-APR-1999; 99US-0128234P.
 PR 08-APR-1999; 99US-0128714P.
 PR 16-APR-1999; 99US-0129845P.
 PR 19-APR-1999; 99US-0130077P.
 PR 21-APR-1999; 99US-0130449P.
 PR 23-APR-1999; 99US-0130510P.
 PR 23-APR-1999; 99US-0130891P.
 PR 28-APR-1999; 99US-0131449P.
 PR 30-APR-1999; 99US-0132048P.
 PR 30-APR-1999; 99US-0132407P.
 PR 04-MAY-1999; 99US-0132484P.
 PR 05-MAY-1999; 99US-0132485P.
 PR 06-MAY-1999; 99US-0132486P.
 PR 06-MAY-1999; 99US-0132487P.
 PR 07-MAY-1999; 99US-0132863P.
 PR 11-MAY-1999; 99US-0134256P.
 PR 14-MAY-1999; 99US-0134218P.
 PR 14-MAY-1999; 99US-0134219P.
 PR 14-MAY-1999; 99US-0134221P.
 PR 14-MAY-1999; 99US-0134370P.
 PR 18-MAY-1999; 99US-0134768P.
 PR 19-MAY-1999; 99US-0134941P.
 PR 20-MAY-1999; 99US-0135124P.
 PR 21-MAY-1999; 99US-0135353P.
 PR 24-MAY-1999; 99US-0135629P.
 PR 25-MAY-1999; 99US-0136021P.
 PR 27-MAY-1999; 99US-0136392P.
 PR 28-MAY-1999; 99US-0136782P.
 PR 01-JUN-1999; 99US-0137222P.
 PR 03-JUN-1999; 99US-0137528P.
 PR 04-JUN-1999; 99US-0137502P.
 PR 07-JUN-1999; 99US-0137724P.
 PR 08-JUN-1999; 99US-0138094P.
 PR 10-JUN-1999; 99US-0138540P.
 PR 10-JUN-1999; 99US-0138847P.
 PR 14-JUN-1999; 99US-0139119P.
 PR 16-JUN-1999; 99US-0139452P.
 PR 16-JUN-1999; 99US-0139453P.
 PR 17-JUN-1999; 99US-0139492P.
 PR 18-JUN-1999; 99US-0139454P.
 PR 18-JUN-1999; 99US-0139455P.
 PR 18-JUN-1999; 99US-0139456P.
 PR 18-JUN-1999; 99US-0139457P.
 PR 18-JUN-1999; 99US-0139458P.
 PR 18-JUN-1999; 99US-0139459P.
 PR 18-JUN-1999; 99US-0139460P.
 PR 18-JUN-1999; 99US-0139461P.
 PR 18-JUN-1999; 99US-0139462P.
 PR 18-JUN-1999; 99US-0139463P.
 PR 18-JUN-1999; 99US-0139750P.
 PR 18-JUN-1999; 99US-0139763P.
 PR 21-JUN-1999; 99US-0139817P.
 PR 22-JUN-1999; 99US-0139899P.
 PR 23-JUN-1999; 99US-0140353P.
 PR 23-JUN-1999; 99US-0140354P.
 PR 24-JUN-1999; 99US-0140695P.
 PR 28-JUN-1999; 99US-0140823P.
 PR 29-JUN-1999; 99US-0140991P.
 PR 30-JUN-1999; 99US-0141287P.
 PR 01-JUL-1999; 99US-0141842P.
 PR 01-JUL-1999; 99US-0142154P.
 PR 02-JUL-1999; 99US-0142055P.
 PR 06-JUL-1999; 99US-0142390P.
 PR 08-JUL-1999; 99US-0142803P.
 PR 09-JUL-1999; 99US-0142920P.
 PR 12-JUL-1999; 99US-0142977P.
 PR 13-JUL-1999; 99US-0143542P.
 PR 14-JUL-1999; 99US-0143624P.
 PR 15-JUL-1999; 99US-0144005P.
 PR 16-JUL-1999; 99US-0144085P.
 PR 16-JUL-1999; 99US-0144086P.
 PR 19-JUL-1999; 99US-0144325P.
 PR 19-JUL-1999; 99US-0144331P.
 PR 19-JUL-1999; 99US-0144332P.
 PR 19-JUL-1999; 99US-0144333P.
 PR 19-JUL-1999; 99US-0144334P.
 PR 19-JUL-1999; 99US-0144335P.
 PR 20-JUL-1999; 99US-0144352P.
 PR 20-JUL-1999; 99US-0144632P.
 PR 20-JUL-1999; 99US-0144684P.
 PR 21-JUL-1999; 99US-0144814P.
 PR 21-JUL-1999; 99US-0145086P.
 PR 21-JUL-1999; 99US-0145088P.
 PR 22-JUL-1999; 99US-0145085P.
 PR 22-JUL-1999; 99US-0145087P.
 PR 22-JUL-1999; 99US-0145089P.
 PR 22-JUL-1999; 99US-0145192P.
 PR 23-JUL-1999; 99US-0145145P.
 PR 23-JUL-1999; 99US-0145218P.
 PR 23-JUL-1999; 99US-0145224P.
 PR 26-JUL-1999; 99US-0145276P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 27-JUL-1999; 99US-0145918P.
 PR 27-JUL-1999; 99US-0145919P.
 PR 28-JUL-1999; 99US-0145951P.
 PR 02-AUG-1999; 99US-0146386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 02-AUG-1999; 99US-0146389P.
 PR 03-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 04-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147192P.
 PR 05-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.


```

PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 28-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 28-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 8,41e-53
Score: 674.50
Percent Similarity: 65.36%
Best Local Similarity: 50.36%
Query Match: 44.35%
DB: 3

US-10-042-894A-8 (1-289) x AAC33685 (1-1104)

QY 4 LeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyPro 23
Db 132 CTCAAGTCCTGACATCAGTTGTCAGGACATTCCTAAACGGGAGCTGTCT 191

QY 24 LeuIleAspGlySerGlyLeuPheTyrLeuProLeuGlnAlaGlyAspArgGlyGluHis 43
Db 192 CTCGTAGATCAAGGTCGGTTCTTCAAGCCACTTCAGGGCGATTCCTCGTGTGAATC 251

QY 44 GluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgIleArgAsp 63
Db 252 GAGGTAAAGTTCTACGAATCTTTCTCTCAACACACAGAGGTTCCAGAACACATCCAT--- 308

```

```

QY 64 ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83
Db 309 AGATATTTCCCGTGTATCAGCGCACTCAAGCAGTT-----CAAGGTTCTGATGAGCA 362

QY 84 ProHisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaProCysValAla 103
Db 363 -----GCCATCATGTGTGGAAAAATCTTTCAGAAATACTCAAAACCATCAGTAATG 416

QY 104 AspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyrIleAlaLys 123
Db 417 GATGTTAAGATGGTTTCGAGAAACATGTCATCTCTGATCTGAGAAATACATCCAAAAA 476

QY 124 CysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyVal 143
Db 477 TGTTTGAAGAAGACACGGGTACCAACACCGTGTTCATCGGGTTTCAGGATCTCTGGTTC 536

QY 144 ArgValValGly---ProGluGlyAlaValTrpArgThrGluArgProGluValLysAla 162
Db 537 GAAGTGTATGATCACAAGAAATCGAGTTTCTGGAAGCCCGAGAGGAAGCTTCTTCGCGG 596

QY 163 MetAspThrAlaGlyValArgValLeuArgValTyrValSer-----SerValAla 180
Db 597 CTCGATGTAGTGGACGCGAGATTGACTCTGAGAGAGTTTGTATCATCTAACTCATTTCG 656

QY 181 AspGluGlyMet-----AspCysAlaLeuAlaAlaValTyrGlyGlyLysGly 198
Db 657 GACACTGGCTCGAAACCTGACTCTGCTTTGCTCGAGTGTTCACGCGGTTCGCCACGGG 716

QY 199 ValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPhe 218
Db 717 ATCTTAACGCAGTGTCTGGAACCTCAAGACCTGGTTCGAGAACCAACGCTCTACCATTC 776

QY 219 TyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGly 238
Db 777 AACTCTTGTTCGATTTTAATGCTCTATGAGATGATCCATCTTGNAGGGAATGATGAT 836

QY 239 GlyGly---ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyVal 257
Db 837 GATGCTAGACCACCAAGTCAAGCTGGTGGATTTGCTCATGTTCTTGATGTTAATGTTGC 896

QY 258 IleAspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleVal 277
Db 897 ATTGACCAATACTTCTTGGTGGTCTTTCCTCTTCTTATCAATAACTTCATCTGAGATTCT 956

RESULT 13
AAC39023
ID AAC39023 standard; DNA; 1168 BP.
XX AC AAC39023;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 23077.
XX KW Hybridisation assay; genetic mapping; gene expression control;
XX KW protein identification; signal transduction pathway; metabolic pathway;
XX KW promoter; termination sequence; ss.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
XX PR 05-MAR-1999; 99US-0123180P.
XX PR 09-MAR-1999; 99US-0123548P.
XX PR 23-MAR-1999; 99US-0125788P.
XX PR 25-MAR-1999; 99US-0126264P.
XX PR 29-MAR-1999; 99US-0126785P.
XX PR 01-APR-1999; 99US-0127462P.

```

PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0134449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.

PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149923P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153700P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-015753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.

PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149829P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158023P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.

PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.: 9,74e-53 Length: 1243
Score: 674.50 Matches: 141
Percent Similarity: 65.36% Conservatives: 42
Best Local Similarity: 50.34% Mismatches: 86
Query Match: 44.35% Indels: 11
DB: 3 Gaps: 7

US-10-042-894A-8 (1-289) x AAC38692 (1-1243)

QY 4 LeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyPro 23
Db 269 CTCAAAGTCCCTGAACATCAGGTTCAGGACACATGCTAAAGCGGAAGCTGCTCT 328
QY 24 LeuIleAspGlySerGlyLeuPheTyrProLeuGlnAlaGlyAspArgGlyGluHis 43
Db 329 CTCGTAGATGACAAAGGTCGGTTCCTCAAGCCACTTCAGGGGATTCCTGCTGTAATC 388
QY 44 GluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArgIleArgasp 63
Db 389 GAGGTAAGTTCTACGAATCTTCTCCTCAACACAGAGGTTCCAGAACACATCCAT 445
QY 64 ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83
Db 446 AGATATTTCCCGTGTATCCGGCACTCAAGCAGTT-----GAAGTTCTCATGAGCA 499
QY 84 ProHisProHisValLeuAspLeuLeuAlaGlyPheGlnAlaProCysValAla 103
Db 500 -----GCCATCATGCTGTGAAATCTTTCAGAAATACCTCAAAACCATCATGTAATG 553
QY 104 AspIleLysIleGlyAlaIleThrProProSerSerProGluProTyrIleAlaLys 123
Db 554 GATGTTAAGTGGTTCGAGAACATCGATTCCTGATCGATCTGGAAGATACATCCAAAA 613
QY 124 CysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyVal 143
Db 614 TGTTTGAGAAAGACACGGGTACCAACACCGGTGTCATCGGTTTCAGATCTCTGTTTC 673
QY 144 ArgValValGly-----ProGluGlyAlaValThrArgThrGluArgProGluValLysAla 162
Db 674 GAAAGTATGATCACAAAGAACATCGAGTTTTCGAAAGCCGAGAGGAAGTCTTCGCGGG 733
QY 163 MetAspThrAlaGlyValArgArgValLeuArgArgTyrValSer-----SerValAla 180
Db 734 CTCGATGTAGTGGAGCGGATGATGACTCTGAGGAAGTTTGATCATCTAACTCATCTTCG 793
QY 181 AspGluGlyMet-----AspCysAlaLeuAlaAlaValTyrGlyGlyLysGly 198
Db 794 GACACTGGCTCGAAACCTGACTCTGCTTTGCTCGAGTGTTCACGGGTTTCCCGCGG 853
QY 199 ValLeuSerGlnLeuArgGluLeuLysAlaThrPheGluGlnThrLeuPheHisPhe 218
Db 854 ATCTTAACGCACTGCTGGAACCTCAAGACCTGCTGAGAACCAACCGCTTACCATTC 913
QY 219 TyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGly 238
Db 914 AACTCTGTGTGATTTTAAGTCTATGAGATGATTCATCTTGAAGGAATGATGAT 973
QY 239 GlyGly-----ValThrValLysLeuValAspPheAlaHisValAlaGlyAspGlyVal 257
Db 974 GATGCTAGACCACAAAGTCAAGCTGGTGGATTTTGCTCATGTTCTTGTGTAATGCTGTC 1033
QY 258 IleAspHisAspPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleVal 277
Db 1034 ATTGACATAACTTCTTGGGTGGTCTTCTCTCTTTCATATAAATTCATCTGATGATCTT 1093

RESULT 15
AAD43515

ID AAD43515 standard; DNA; 1105 BP.

XX AC AAD43515;

XX 14-NOV-2002 (first entry)

XX Soybean inositol polyphosphate kinase (IPPK) DNA.

XX Soybean; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
XX nutritional value; animal feed; transgenic; gene; ds.

XX Glycine max.

XX Key Location/Qualifiers
XX CDS 12..851
XX /*tag= a
XX /product= "Soybean IPPK protein"

XX WO200259324-A2.

XX 01-AUG-2002.

XX 09-JAN-2002; 2002WO-US003120.

XX 12-JAN-2001; 2001US-0261465P.

XX (PION-) PIONEER HI-BRED INT INC.

XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;

XX WPI; 2002-636540/68.

XX P-PSDB; AAE26197.

XX New inositol polyphosphate kinase polynucleotides and polypeptides,
XX useful in modulating phytic acid biosynthesis by decreasing phytate or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed.

XX Claim 1; Page 66-67; 86pp; English.

XX The invention relates to novel inositol polyphosphate kinase (IPPK)
XX polypeptides and polynucleotides. Sequences of the invention are useful
XX in modulating the phytic acid biosynthesis by decreasing phytate and/or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed, or to reduce the environmental impact of animal waste.
XX Polynucleotides of the invention are used to produce transgenic plants with an
XX altered phenotype. IPPK proteins are used to screen compounds that
XX modulate their activity and raising anti-idiotypic antibodies. The
XX present sequence is soybean IPPK DNA

XX SQ Sequence 1105 BP; 279 A; 311 C; 233 G; 282 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.94e-50 Length: 1105
Score: 644.50 Matches: 141
Percent Similarity: 62.72% Conservatives: 34
Best Local Similarity: 50.54% Mismatches: 85
Query Match: 42.37% Indels: 19
DB: 6 Gaps: 8

US-10-042-894A-8 (1-289) x AAD43515 (1-1105)

QY 7 ProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyProLeuIleAsp 26

Db 24 CCGAGACACCAAGTGGCGGCGCAACAGGCCAAGCAATCTCTGGGCCCTCGTCGAC 83

QY 27 GlySerGlyLeuPheTyrLysProLeuGlnAlaGly-----AspArgGlyGlu 42

Db 84 GATTTTGAAAATTTCTACAGCCCTCCAGACCCCTCCAGACCAACAAAGACGACGACCCCGGCTCC 143

QY 43 HisGluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArgIleArg 62

Db 144 ACCGAACTCTCTTTTACCTCTCTCGCGCC---GCGGCCACGACTACTCTCATCGCG 200

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 27, 2004, 04:39:30 ; Search time 4448 Seconds
(without alignments)
2816.128 Million cell updates/sec

Title: US-10-042-894A-8

Perfect score: 1521

Sequence: 1 MSDLHPPEHQVAGHRASASK.....IKFVSDIVPETPTQPLGPS 289

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2 1/USPTO.spool/US10042894/runat 24032004 145047 16946/app query.fasta_1.455
-DB=GenEmbl -OPMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LIST=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=humar40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=spt -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10042894.cgn 1 1 3731 @runat 24032004 145047 16946 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl :
1: gb ba : *
2: gb htg : *
3: gb in : *
4: gb om : *
5: gb ov : *
6: gb pat : *
7: gb ph : *
8: gb pl : *
9: gb pr : *
10: gb ro : *
11: gb sts : *
12: gb sy : *
13: gb un : *
14: gb vl : *
15: em ba : *
16: em fun : *
17: em hum : *
18: em in : *
19: em mu : *
20: em om : *
21: em or : *
22: em ov : *
23: em pat : *
24: em ph : *
25: em pl : *
26: em ro : *
27: em sts : *
28: em un : *

29: em vi : *
30: em htg hum : *
31: em htg inv : *
32: em htg other : *
33: em htg mus : *
34: em htg pln : *
35: em htg rod : *
36: em htg mam : *
37: em htg vrt : *
38: em sy : *
39: em htg hum : *
40: em htg mus : *
41: em htg other : *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1521	100.0	1344	6	AX513570 Sequence
2	1485	97.6	923	6	AX513568 Sequence
3	1435	94.3	923	6	AX513566 Sequence
4	1406.5	92.5	1169	6	AX513564 Sequence
5	1307	85.9	3416	6	AX513583 Sequence
6	1233	81.1	899	6	AX513578 Sequence
7	1148.5	75.5	132170	8	AP005749 Oryza sat
8	1144.5	75.2	1570	8	AK072296 Oryza sat
9	808.5	53.2	643	6	AX513580 Sequence
10	742.5	48.8	1020	6	AX513576 Sequence
11	722.5	47.5	1195	6	AX513574 Sequence
12	718.5	47.2	25054	8	LUS310150 Linum usi
13	714.5	47.0	903	8	ATH243592 Arabidops
14	714.5	47.0	903	8	AY072621 Arabidops
15	714.5	47.0	903	8	AY147936 Arabidops
16	714.5	47.0	1166	8	AF412073 Arabidops
17	714.5	47.0	1170	8	ATH245521 Arabidops
18	714.5	47.0	57246	8	AB010069 Arabidops
19	704.5	46.3	1130	8	AY085862 Arabidops
20	674.5	44.3	1243	8	AY087217 Arabidops
21	673.5	44.3	1390	8	AY136378 Arabidops
22	673.5	44.3	116763	8	ATT211
23	664	43.7	115040	8	AC124954 Medicago
24	658.5	43.3	861	8	ATH404678 Arabidops
25	658.5	43.3	861	8	AY147935 Arabidops
26	648	42.6	129248	2	AC127018 Medicago
27	644.5	42.4	1105	6	AX513572 Sequence
28	469.5	30.9	43428	8	AP004981 Lotus cor
29	464	30.5	519	6	AX513581 Sequence
30	340.5	22.4	116893	8	AP004772 Oryza sat
31	329	21.6	353	6	AX513582 Sequence
32	288	18.9	85095	3	AC004573 Drosophil
33	288	18.9	142257	2	AC017491 Drosophil
34	288	18.9	182726	3	AC008002 Drosophil
35	288	18.9	307741	3	AE003589 Drosophil
36	286	18.8	1219	3	AY119207 Drosophil
37	286	18.8	1219	3	AY119207 Drosophil
38	217.5	14.3	3003	3	AK116726 Ciona int
39	201.5	13.2	1251	9	AF432853 Homo sapi
40	199.5	13.1	5432	10	BC052463 Mus muscu
41	194	12.8	5432	10	AY014898 Rattus no
42	193	12.7	1341	6	AR380264 Sequence
43	192	12.6	2391	6	AX816462 Sequence
44	192	12.6	2391	9	AF393812 Homo sapi
45	192	12.6	2890	6	AX504248 Sequence

ALIGNMENTS

RESULT 1

```

AX513570
LOCUS AX513570 1344 bp DNA linear PAT 05-OCT-2002
DEFINITION Sequence 7 from Patent WO02059324.
ACCESSION AX513570
VERSION AX513570.1 GI:23559670
KEYWORDS
SOURCE
ORGANISM
Zeal mayas
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.
Novel inositol polyphosphate kinase genes and uses thereof
Patent: WO 02059324-A 7 01-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source
Location/Qualifiers
1..1344
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
52..921
/feature="unnamed protein product"
/codon_start=1
/protein_id="CAD52954.1"
/db_xref="GI:23559671"
/db_xref="REMBL:CAD52954"
/translation="MSDLHPPEHQVAGHRASAKLPLIDGSLFYKPLQAGDRGEHE
VAFYFAFAHAAPARIDRTFFPHGHRLLPTEAPQGPPEPHLVLDLLAGFQAPCV
ADIKGATITWPPSSPEPIAKLMDRGTTVLLGFRVSVGVVPEGAVWRTERPEV
KAMDAGVRVLRVSVVADGMDALAAAVYGGKGVLSQLRELKAWFEQTLFHF
YSASILLGYDAAVAAGDGGGVTVKLVDFAHVAGDGVIDHNFGLGSLIKFVSDI
VPETPTQPLGPS"
ORIGIN
Alignment Scores:
Pred. No.: 6.98e-106 Length: 1344
Score: 1521.00 Matches: 289
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-042-894A-8 (1-289) x AX513570 (1-1344)
QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20
DB 52 ATGTCGACCTCCACCGCGGAGCACCAGTCCGCGGCCACCGCGCTCCGCGCAGCAAG 111
QY 21 LeuGlyProLeuLeuAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
DB 112 CTGGGCGCGCTCATCGACGCTCCGCGCTCTTCTTCAAGCCGCTCCAGCGCGCGACCGT 171
QY 41 GlyGluHisGlnValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
DB 172 GGGAGCAGAGTCCGCTTCTATGAGGAGTCTCCGCGCCACCGCGCTCCGCGCGC 231
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
DB 232 ATCCGAGACACCTTCTTCCCGCGGTTCACCGCACCGAGCTCTCCACCGCGCGCAG 291
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100
DB 292 CCGGGAGCGGATCCTCCTCCTGCTGCTGAGGACCTCTCTCGGGGTTTCAGGCGCC 351
QY 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSerProGluProTyr 120
DB 352 TGGCTCGCAGACATCAAGTCGGCGCATCATCGTGGCCACCGAGTTTCGCGGAGCCCTAC 411
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
DB 412 ATGCCAAGTGCCTCGCCAGGACCGCGGACCCAGGCGTTCTGCTCGGATTCGCGCTC 471

```

```

QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
DB 472 TCCGCGCTCCGAGTCTCGGCGCCGAGGCGCGCTGTGGCGAGCGGAGCGCGGAGTG 531
QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgfyrValSerSerValAla 180
DB 532 AAGGCCATGGACACCGCGCGCTCCGCGCGTGTCTCCGCGCTACGTCTCCGCTGCC 591
QY 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTrpGlyGlyGlyGlyValLeu 200
DB 592 GACGAGGGGATGGACTGTGCGCTCCGCGCGCTGTACGAGGAAAGGAGGAGTCTTG 651
QY 201 SerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSer 220
DB 652 TCACAGCTCGCGAGCTCAAGCGTGGTTCGAGGAGCAGACTCTGTCCACTTCTACTCG 711
QY 221 AlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyAspGlyGly 240
DB 712 GCGTCGATTTCTTCTGGGCTATGATGCTGCTCAGTCGCGAGCGGAGATGGGGTGG 771
QY 241 ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHis 260
DB 772 GTGACGGTGAAGCTGGTGGACTTTGCCCATGTGCCCGAGGTGATGGGTGATTCACCC 831
QY 261 AsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluThr 280
DB 832 AACTTCTCGCGCGGCTCTGCTCGCTGATCAAGTTCGTTCTGACATTTGTCGAGACT 891
QY 281 ProHisThrGlnProLeuGlyProSer 289
DB 892 CCTCATACGACGCTTTGGGTCTTCT 918
RESULT 2
AX513568
LOCUS AX513568 923 bp DNA linear PAT 05-OCT-2002
DEFINITION Sequence 5 from Patent WO02059324.
ACCESSION AX513568
VERSION AX513568.1 GI:23559668
KEYWORDS
SOURCE
ORGANISM
Zeal mayas
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.
Novel inositol polyphosphate kinase genes and uses thereof
Patent: WO 02059324-A 5 01-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source
Location/Qualifiers
1..923
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
53..922
/feature="unnamed protein product"
/codon_start=1
/protein_id="CAD52953.1"
/db_xref="GI:23559669"
/translation="MPDLHPPEHQVAGHRASAKPLIDGSLFYKPLQAGDRGEHE
VAFYFAFAHAAPARIDRTFFPHGHRLLPTEAPQGPPEPHLVLDLLAGFQAPCV
ADIKGATITWPPSSPEPIAKLMDRGTTVLLGFRVSVGVVPEGAVWRTERPEV
KAMDAGVRVLRVSVVADGMDALAAAVYGGKGVLSQLRELKAWFEQTLFHF
YSASILLGYDAAVAAGDGGGVTVKLVDFAHVAGDGVIDHNFGLGSLIKFVSDI
VPETPTQPLGPS"
ORIGIN
Alignment Scores:
Pred. No.: 2.35e-103 Length: 923
Score: 1485.00 Matches: 283
Percent Similarity: 98.27% Conservative: 1
Best Local Similarity: 97.92% Mismatches: 5

```

```
Query Match: 97.63% Indels: 0
DB: 6 Gaps: 0
US-10-042-894A-8 (1-289) x AX513568 (1-923)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerHisSerHis 20
DB 53 ATGCCGCGACCTCCACCGCGGAGACCAAGATCGCGGTACCGCGCTCCGCCAGCAG 112
QY 21 LeuGlyProLeuLeuAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
DB 113 CCGGCGCGCTCATCGACGCTCCGGCTCTTCTACAGCCGCTCCAGCGCGCGACCGT 172
QY 41 GlyGluHisGlnValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
DB 173 GGGAGACGAGTTCCTTCTATGAGCGTTCCTCGGCCACGCGCGCTCCGCCCGCGC 232
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
DB 233 ATCCGAGACACCTTCTCCCGCGGTTCACGCGCACGCGACTCTCCACCGCGCGCAG 292
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100
DB 293 CCGGCGGAGCGCATCCGCACTCTCTCTCGACGACTCTCTCGCGGATTTGAGCGCGCC 352
QY 101 CysValAlaAspIleLeuGlyProGluGlyAlaValTyrArgThrGluArgProGluProTyr 120
DB 353 TCGGTCCGAGACATCAAGATCGCGCCATCAGTGGCCACCGAGTTCGCGGAGCCCTAC 412
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
DB 413 ATCCGCAAGTGCCTCGCATGACCGCGGACCGACGAGCTCTCTCGGATTCGCGCTC 472
QY 141 SerGlyValArgValGlyProGluGlyAlaValTyrArgThrGluArgProGluVal 160
DB 473 TCCGCGTCCGAGTCTCGGCCCGCGCGCGCTGTGGCGGACGCGCGCGGAGGTG 532
QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgTyrValSerSerValAla 180
DB 533 AAGGCGATGAGACACCGCGCGGTCCGCGCGTGTCTCCGCGTACGTCTATCGGTGCC 592
QY 181 AspGluGlyMetAspCysAlaLeuAlaValAlaValTyrGlyLysGlyGlyValLeu 200
DB 593 GACGAGGGGATGGATGTGCGCTCGCGCGCGGTGTACGAGGAGAAAGGTGGAGTCTTG 652
QY 201 SerGlnLeuArgGlyLeuLysAlaTyrPheGluGluGlnThrLeuPheHisPheTyrSer 220
DB 653 TCACAGCTCGCGAGCTCAAGCGGTGTTCGAGGAGCAGACTCTGTCTTCTACTCG 712
QY 221 AlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGly 240
DB 713 GCGTGCATCTTCTGGGCTATGATCTGCTGCACTCGAGTTCGAGCGAGCGGAGGTGG 772
QY 241 ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHis 260
DB 773 GTGACGGTGAAGCTGTGGACTTTCCCATGTCGCGCGGAGGTGAAGGGGTGATTGACCAC 832
QY 261 AsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluThr 280
DB 833 AACTCTCTGGGCGGCTCTGCTCGCTGATCAAGTCTGTTCTGACATTTGTCAGAGACT 892
QY 281 ProHisThrGlnProLeuGlyProSer 289
DB 893 CCTCAGACGCGCTTTGGGTCTCTTCT 919

RESULT 3
AX513566
LOCUS AX513566 923 bp DNA linear PAT 05-OCT-2002
DEFINITION Sequence 3 from Patent WO02059324.
ACCESSION AX513566
VERSION AX513566.1 GI:23559666
KEYWORDS Zea mays
```

```
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Caboon, R.E.
TITLE Novel inositol polyphosphate kinase genes and uses thereof
JOURNAL Patent: WO 02059324-A 3 01-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
    1..923
    /organism="Zea mays"
    /mol_type="unassigned DNA"
    /db_xref="taxon:4577"
    53..736
    /note="unnamed protein product"
    /codon_start=1
    /protein_id="CAD52952.1"
    /db_xref="GI:23559667"
    /translation="MPDLHPPEHGVAGHRASAKPGLIDGSLFYKPLQAGDRGEHE
    VAFYEATSAHAAPARIDTFFPHGTRLLPTEAQCPEPHPLVLDLLAGFEAPCV
    ADIKGALTWPSPSPPEYIAKAMDSTTSVLLGERVSGVYVVEGAVWRTPEPV
    KAMDTGVRRVLRIRYVSACRRGDLRARGGVRRKWSLVTAARAQGVVRGAASVPL
    LLGVDSGL"

ORIGIN
Alignment Scores:
Pred. No.: 1..4e-99 Length: 923
Score: 1435.00 Matches: 277
Percent Similarity: 96.21% Conservative: 2
Best Local Similarity: 95.52% Mismatches: 10
Query Match: 94.35% Indels: 1
DB: 6 Gaps: 0

US-10-042-894A-8 (1-289) x AX513566 (1-923)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerHisSerHis 20
DB 53 ATGCCGCGACCTCCACCGCGGAGACCAAGATCGCGGTACCGCGCTCCGCCAGCAG 112
QY 21 LeuGlyProLeuLeuAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
DB 113 CCGGCGCGCTCATCGACGCTCCGGCTCTTCTACAGCGCGCTCCAGCGCGCGACCGT 172
QY 41 GlyGluHisGlnValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
DB 173 GGGAGACGAGGTTCCTTCTATGAGCGGTTCCTCGGCCACGCGCGCTCCGCCCGCGC 232
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
DB 233 ATCCGAGACACCTTCTCCCGCGGTTCACGCGCACGCGACTCTCTCCACCGAGCGCAG 292
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100
DB 293 CCGGCGGAGCGCATCCGCACTCTCTCTCGACGACTCTCTCGCGGATTTGAGCGCGCC 352
QY 101 CysValAlaAspIleLysIleGlyAlaValTyrProProSerSerProGluProTyr 120
DB 353 TCGGTCCGAGACATCAAGATCGCGCGCTCATCGTGGCCACCGAGTTCGCGGAGCCCTAC 412
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
DB 413 ATCCGCAAGTGCCTCGCATGACCGCGGACCGACGAGCTCTCTCGGATTCGCGCTC 472
QY 141 SerGlyValArgValGlyProGluGlyAlaValTyrArgThrGluArgProGluVal 160
DB 473 TCCGCGTCCGAGTCTCGGCCCGCGCGCTGTGGCGGACGCGCGCGGAGGTG 532
QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgTyrValSerSerValAla 180
DB 533 AAGGCTATGAGACACCGTCCGCGGTCTCGCGCGTGTCTCGCGCGCTACGTCTATCGCTTGC 592
```

```
Qy 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyTysGlyGlyValLe 200
Db 593 CGACGAGGGATGACTCGCGCTCGCGCGCGGTGTACGAGAGAAAGGTGGAGTCT 652
Qy 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220
Db 653 GTCACTGTGCGCGAGCTCAAGCGGTGTTCGAGGAGCAGCTCTGTTCCTTCACTTCTACTC 712
Qy 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGlyVal 240
Db 713 GCGCGTGCATTTCTTGGGCTGATGCTCTGCTGAGTTCGAGCGAGCGAGGTGGGGTGG 772
Qy 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260
Db 773 GGTAAACAGTGAAGCTGTGTGACTTTGCCATGTGCGCGAGGATGATGGGTGATTTGACCA 832
Qy 260 sAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280
Db 833 CAACCTTCCTGCGCGGCTCTGTCTAGTATCAAGTTCGTTTCTGACATTTGTCAGATTTCCAGAGAC 892
Qy 280 rProHisThrGlnProLeuGlyProSer 289
Db 893 TCCTCAGACGCGCTTTGGGTCTCTCT 920

RESULT 4
AX513564 1169 bp DNA linear PAT 05-OCT-2002
LOCUS
DEFINITION
ACCESSION AX513564
VERSION AX513564.1 GI:23559664
KEYWORDS
SOURCE
ORGANISM
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
1
REFERENCE
Shi,J., Beach,L.R., Wang,H., Rafalski,J.A. and Cahoon,R.E.
AUTHORS
Novel inositol polyphosphate kinase genes and uses thereof
TITLE
Patent: WO 02059324-A 1 01-AUG-2002;
JOURNAL
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
Location/Qualifiers
1..1169
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
84..806
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD52951.1"
/db_xref="GI:23559665"
/db_xref="REMBL:CAD52951"
/translation="WPDLLPPEHVOVAGHRSASAKLGLIDGSLFKYKPLQAGDRGEHE
VAFYERASHAHPARIRITFFRPHGTLLTPEAQCPHPLHLVLDLLAGFEAPCV
ADIKIGAITWSPSPPEYIAKIAKDRGTSLVLLGRVLRSPRRGRKRVADGAGGE
GTHRRRPRAALRVIRPRTGVTGRTARSRRCTEKEVSCASSRHGRSLCSTS
TRRRFPWAMLLQSQOAEVGVG"

CDS
Alignment Scores:
Pred. No.: 2,61e-97 Length: 1169
Score: 1406.50 Matches: 277
Percent Similarity: 96.21% Conservative: 2
Best Local Similarity: 95.52% Mismatches: 10
Query Match: 92.47% Indels: 2
DB: 6 Gaps: 0

ORIGIN
US-10-042-894A-8 (1-289) x AX513564 (1-1169)
Qy 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20
Db 84 ATGCCCGACCTCCACCCCGCGAGCACCAAGTCGCCGCTCACCGCGCTCCGCGAGCAG 143
```

```
Qy 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrIrisProLeuGlnAlaGlyAspArg 40
Db 144 CTGGGCGCGCTCTATCGACGGCTCCGGCTCTCTTACAAAGCGCGTCCAGGCGCGGACCGT 203
Qy 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArg 60
Db 204 GGGAGGACAGAGTTCGCTCTATGAGCGGTCTCCGCCACGCGCGCTCCGCGCGCGC 263
Qy 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 264 ATCCGAGACACCTTCTTCCCGCGGTTCACCGCACCGGACTCTCTCCACCGAGCGGAG 323
Qy 81 ProGlyGluProHisProHisLeuValLeuAspPheLeuAlaGlyPheGlnAlaPro 100
Db 324 CCCGGGAGCGCATTCGCGACCTCGTCTCGACACCTCTCTCGCGGGTCTTGAGCGCGCC 383
Qy 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyr 120
Db 384 TCGTTCGACAGACATCAAGATCGGCGCCATCATCGTGGCCACCGAGTTCGCGAGCCCTAC 443
Qy 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 444 ATGCCAAGTACCTCGCCAGGACCGCGGGACCAAGCGGTCTCTCGGATTCGCGCTC 503
Qy 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
Db 504 TTGC--GTCCAGTGTGTCGGCCCGAGGCGCGGTGTGGCGAGCGAGCGCCCGAGGTG 561
Qy 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgTyrValSerSer-ValAla 180
Db 562 AAGCTATGGACACCGTCGGCGTCCGCGCGTCTCGCGCGCTAGTGTCTATCGCTTGC 621
Qy 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyValLe 200
Db 622 CGACGAGGGATGACTCGCGCTCGCGCGCGGTGTACGAGAGAAAGGTGGAGTCTT 681
Qy 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220
Db 682 GTCAAGTGTGCGCGAGCTCAAGGCGATGTTGGAGGAGCAGACTCTGTTCCTTCTTACTC 741
Qy 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGlyVal 240
Db 742 GCGCTCGATTTCTTGGGCTATGATGCTGCTGCAGTCCGACGAGCGAGGTGGGGTGG 801
Qy 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260
Db 802 GGTAAACAGTGAAGCTGTGTGACTTTTCCCATGTGCGCGAGGCGGTGATGGGTGATTTGACCA 861
Qy 260 sAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280
Db 862 CAACCTTCCTGCGCGAGCTCTGTCTAGTATCAAGTTCGTTTCTGACATTTGTTTCCAGAGAC 921
Qy 280 rProHisThrGlnProLeuGlyProSer 289
Db 922 TCCTTAGACGACGCTTTGGGTCTCTCT 949

RESULT 5
AX513583 3416 bp DNA linear PAT 05-OCT-2002
LOCUS
DEFINITION
ACCESSION AX513583
VERSION AX513583.1 GI:23559684
KEYWORDS
SOURCE
ORGANISM
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
1
REFERENCE
Shi,J., Beach,L.R., Wang,H., Rafalski,J.A. and Cahoon,R.E.
AUTHORS
Novel inositol polyphosphate kinase genes and uses thereof
TITLE
Patent: WO 02059324-A 20 01-AUG-2002;
JOURNAL
PIONEER HI-BRED INTERNATIONAL, INC. (US)
```



```

FEATURES             Location/Qualifiers
     source            1..3416
                        /organism="Zea mays"
                        /mol_type="unassigned DNA"
                        /db_xref="taxon:4577"
     CDS               72..407
                        /note="unnamed protein product"
                        /codon_start=1
                        /protein_id="CAD52959.1"
                        /db_xref="GI:23559685"
                        /translation="MPDLHPPEHQVAGHRASAKLGPLIDSGLYFKPLOGDRGEHE
                        VAFYEAFSAHAAPARIRDTFFPHFGTLLPTEAQGPPEHPHVLVDLLAGFEAPCV
                        ADIKIGAIT"

ORIGIN
Alignment Scores:
Pred. No.:          2,99e-89          Length:          3416
Score:             1307.00           Matches:          260
Percent Similarity: 90.34%           Conservative:    2
Best Local Similarity: 89.66%         Mismatches:      9
Query Match:       85.93%            Indels:         20
DB:                6                Gaps:            1

US-10-042-894A-8 (1-289) x AX513583 (1-3416)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerIlys 20
Db 72 ATGCCGACCTCCACCCCGCGGAGCACCAGTCGCGGTACCCGCGTCCGCCAGCAAG 131
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrGlyProLeuGlnAlaGlyAspArg 40
Db 132 CTGGGCCCACTATCGAGCATCTGGCCCTCTTACAGCCGCTCCAGCGCGGCGACCGT 191
QY 41 GlyGluHisGlnValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
Db 192 GGGGAGCAGAGTCGCTTCTATGAGCGGTCTCCGCCACCGCGCGTCCCGCGCGC 251
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 252 ATCCGAGACACCTCTTCCCGCGGTTCACGCGCACGCGACTCTCTCCACCGAGCGCGAG 311
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100
Db 312 CCCGGGAGCGCATCCGCCTCTGCTTCAGCAGCTCTCTCGCGGGTTTGAGGCGGCC 371
QY 101 CysValAlaAspIleLysIleGlyValAlaIleThrTyrProProSerSerProGluProTyr 120
Db 372 TGGTCGCAGACATCAAGATCGGTGCCATCAGTG----- 406
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 407 -----ACCGAGCGCTCTGCTCGGATTCCCGGTC 436
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTyrArgThrGluArgProGluVal 160
Db 437 TCCGGCGTCCGAGTCGTGCGGCCCGAGGGCCCGTGTGGCGAGCGAGCGGCCGAGGTG 496
QY 161 LysAlaMetAspThrAlaGlyValArgValArgValLeuArgArgTyrValSerSer-ValAl 180
Db 497 AAGGTATGACATTTGTCGGCGTCCGCGCGTGTCTCGCGCGCTACGTGTCTACGCTTGC 556
QY 180 aAspGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyValLe 200
Db 557 CGACGAGGGATGACTTCGCGGCTCGCGCGCGCGGTGTACGGAGAGAAAGGTGGAGTCT 616
QY 200 userGlnLeuArgGlnLeuLysAlaTyrPheGluGlnThrLeuPheHisPheTyrSe 220
Db 617 GTCAAGCTGCGGAGCTCAAGGCGTGTTCAGGGGCGAGCTCTGTTCACATCTTCTACTC 676
QY 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaGlyGlyAspGlyGly 240
Db 677 GGCCTCGATTCTTCTGGGCTATGATGCTGCTGTCAGTCGACGAGCGAGGTGGGGGTGG 736

```

```

QY 240 YValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260
Db 737 GGTAAACAGTGAAGCTGGTGGACTTTGCCCATGTGCCGAGGCTGATGGGTGATTGACCA 796
QY 260 sAenPheLeuGlyGlyLeuCysSerLeuLeuLeuPheValSerAspIleValProGluTh 280
Db 797 CAACTTCTGGCGGGCTCTGCTAGCTGATCAAGTTTGTCTGACATTGTTCCAGAGAC 856
QY 280 rProHisThrGlnProLeuGlyProSer 289
Db 857 TCCTCAGACGCGCTTTGGTCTCTCT 884

RESULT 6
AX513578
LOCUS             AX513578             899 bp             DNA             linear             PAT 05-OCT-2002
DEFINITION       Sequence 15 from Patent WO02059324.
ACCESSION        AX513578
VERSION          AX513578.1             GI:23559679
KEYWORDS
SOURCE           Zea mays
ORGANISM         Zea mays
REFERENCE
AUTHORS          Shi,J., Beach,L.R., Wang,H., Rafalski,J.A. and Cahoon,R.E.
TITLE            Novel inositol polyphosphate kinase genes and uses thereof
JOURNAL          Patent: WO 02059324-A 15 01-AUG-2002;
                  PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source           1..899
                 /organism="Zea mays"
                 /mol_type="unassigned DNA"
                 /db_xref="taxon:4577"
                 89..424
                 /note="unnamed protein product"
CDS
                 /codon_start=1
                 /protein_id="CAD52958.1"
                 /db_xref="GI:23559680"
                 /db_xref="RENTREMBL:CAD52958"
                 /translation="MPDLHPPEHQVAGHRASAKLGPLIDSGLYFKPLOGDRGEHE
                 VAFYEAFSAHAAPARIRDTFFPHFGTLLPTEAQGPPEHPHVLVDLLAGFEAPCV
                 ADIKIGAIT"

ORIGIN
Alignment Scores:
Pred. No.:          2,42e-84          Length:          899
Score:             1233.00           Matches:          248
Percent Similarity: 88.42%           Conservative:    4
Best Local Similarity: 87.02%         Mismatches:     14
Query Match:       81.07%            Indels:         20
DB:                6                Gaps:            1

US-10-042-894A-8 (1-289) x AX513578 (1-899)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerIlys 20
Db 89 ATGCCGACCTCCACCCCGCGGAGCACCAGTCGCGGTCCCGCGCTCCGCCAGCAAG 148
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrIlysProLeuGlnAlaGlyAspArg 40
Db 149 CTGGGCCCATCTCATCGACCGCTCTGGCCCTCTTACAAAGCCGTCTCCAGCGCGGACCGT 208
QY 41 GlyGluHisGlnValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
Db 209 GGGGAGCAGCGGTGCGCTTCTATGAGCGCTTCTCGCCCCACGCGCGCTCCCGCGCGC 268
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 269 ATCCGAGACACCTTCTTCCCGCGGTTCACGCGCAGCGACTCTCTCCCGCAGCGCGAG 328
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100

```


Db 1163 GTGATTGACCAACTTCTTGGCGGGCTCTGCTCGCTCATCAAGTTTCATCGCGCACATT 1104

Qy 277 ValProGlu 279
||| |||

Db 1103 GTCGCCGAG 1095

RESULT 8
AK072296

LOCUS AK072296 1570 bp mRNA linear PLN 24-JUL-2003

DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J023018G11, full insert sequence.

VERSION AK072296

KEYWORDS FLJ_CDNA; CAP trapper.

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1

AUTHORS The Rice Full-Length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-Length cDNA Project Team:
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group, Ootomo, Y., Murakami, K.,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN:
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Oota, N., Ota, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Yoshino, M., and Hayashizaki, Y.

TITLE Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice

JOURNAL Science 301 (5631), 376-379 (2003)

MEDLINE 22752273

PUBMED 12869764

REFERENCE 2 (bases 1 to 1570)

AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K.,
Numasaki, R., Ohneda, E., Ohtsuki, K., Ooka, M., Ooka, H.,
Oota, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, K., Sakai, K.,
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,
Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and
Yoshimura, A.

Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007)

COMMENT This clone is one of the 28K full-length cDNA clones from japonica
rice.
URL : <http://cdna01.dna.affrc.go.jp/cdna/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,
Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and
Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y.,
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,
Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,
Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,
Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ooka, H., Oota, N.,
Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers
1. 1570
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023018G11"

ORIGIN

Alignment Scores:
Pred No.: 2,23e-77 Length: 1570
Score: 1149.50 Matches: 217
Percent Similarity: 84.10% Conservative: 21
Best Local Similarity: 76.68% Mismatches: 40
Query Match: 75.25% Indels: 5
DB: 8 Gaps: 3

US-10-042-894A-8 (1-289) x AK072296 (1-1570)

Qy 2 SerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeu 21

Db 384 TCCGACCTGGCGCGCGCGGACACAGTGCGGGGACCGCGCGTCCGCGACAGGTG 443

Qy 22 GlyProLeuIleAspGlySerGlyLeuPheTyrllysProLeuGlnAlaGlyAspArgGly 41

Db 444 GGCCCGCTCTGCGCGGAGGGGCTCTTCTACAGGCCCTTCCAGACGGGGAGCGGG 503

Qy 42 GluHisGluValAlaPheTyrlGluAlaPheSerAlaHisAlaValProAlaArgIle 61

Db 504 GAGCACGAGCGCGCTTCTACCGCGCTTCCACCGCGACCGCGCGTCCCGCGGCGTC 563

Qy 62 ArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnPro 81

Db 564 CGGGCGCGCTTCTTCCGCGCTTCCACCGCGACCGCGCTTCTTCCCGGCCCGCCG 623

Qy 82 --GlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro 100

Db 624 GCGCGCGCGCTTCCCGCGACATCTCTCGAGACCTCTCTCGGGGCTTCCCGTCCCC 683

Qy 101 CysValAlaAspIleLysIleGlyValaIleThrProProSerSerProGluProThr 120

Db 684 TGGCTCCGCGAGCTCAAGATCGGCGCTCGACGTTGGCGCGCGGATCCCGGACCCCTAC 743

Qy 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140

Db 744 GTCGCCAAGTGCTCGCGAAGGACCGGAGACCGGAGACCGCGGCTCTCTCGGCTCCG 803

Qy 141 SerGlyValArgValValGlyProGlu--GlyAlaValTrpArgThrGluArgProGlu 159

[illegible]

AX513580	AX513580	643 bp	DNA	linear	PAT 05-OCT-2002
LOCUS	Sequence 17 from Patent WO02059324.				
DEFINITION					
ACCESSION	AX513580				
VERSION	AX513580.1	GI:23559681			

REFERENCE

1. Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.
Novel inositol polyphosphate kinase genes and uses thereof
Patent: WO 02059324-A 17 01-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

AUTHORS

TITLE

JOURNAL

```

ORIGIN
/db_xref="taxon:4577"

Alignment Scores:
  Pred. No.: 1.03e-52
  Score: 808.50
  Length: 643
  Matches: 162
  Conservative: 1
  Mismatches: 10
  Indels: 3
  Gaps: 0
  DB: 6

```

US-10-042-894A-8 (1-289) x AX513580 (1-643)

QY	1	MetSerAspLeuHisProProGluHisGlnValAlaGlyHisArgAlaSerAlaSerIys	20
Db	117	ATCCCGACCTCACCCGCGGAGACCAAGTGC CGCTCACGCGCTCCG CAGCAG	176
QY	21	LeuGlyProLeuLeuAspGlySerGlyLeuPheTyrIysProLeuGlnAlaGlyAspArg	40
Db	177	CTGGGCCCCGTCAFCGACGCTCGGGCTCTTCTACAAGCGCTCCAGGGCGCGACCGT	236

41	QY	GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg	60
237	Db	GGGAGCAGCAGAGTCGCCCTTCATGAGGGGTTCTCCGCCACCGCGTCCCGGCCCGC	296
61	QY	IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln	80
297	Db	ATCCGAGACACCTTCTCCCCCGGTTCACGGCACGCACTCTCTCCCAACCGAGGGCGAG	356
81	QY	ProGlyVgluProHisProHisValLeuAspAspLeuAlaGlyPheGlnAlaPro	100
357	Db	CCCGGGAGCGCATCCGCACCTCGTCTCGACGACCTCTCGCGGGTTGAGGGCGCC	416
101	QY	CysValAlaAspIleIysIleGlyAlaIleThrTyrProProSerSerProGluProTyr	120
417	Db	TGCGTCGCAGACATCAAGATCGCGGCATCACGTGCCACCGAGTTCGCCGGAGCCCTAC	476
121	QY	IleAlaIysCysLeuAlaIysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal	140
477	Db	ATCGNCAAGTACTCTNGCCAGGACCGGGGACACAGAGCGTTCCTGCTCGATTCCGGTTC	536
141	QY	SerGlyValArgValValGlyProGluGlyAlaValTyrArgThrGluArg-ProGluVal	160
537	Db	TTGC--GTCGAGTCGTGCGCCCGAGGCGCGGTGTGCGGACGAGCGCCCGGGGGT	594
160	QY	IlyAlaMetAspThr-AlaGlyValArgArgValLeuArg	173
595	Db	GAANGCTATGGACACCGTCGGNGNCGCGNGTGTCTCGG	635

LOCUS	AX513576	1020 bp	DNA	linear	PAT 05-OCT-2000
SEQUENCE	Sequence 13 from Patent WO02059324.				
ACCESSION	AX513576				
VERSION	AX513576.1	GI:23559677			
KEYWORDS	.				
SOURCE	Parthenium argentatum				
ORGANISM	Parthenium argentatum				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asteraceae; Asteroideae; Heliantheae; Parthenium.				

1. Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E. Novel inositol polyphosphate kinase genes and uses thereof Patent: WO 02059324-A 13 01-AUG-2002; PIONEER HI-BRED INTERNATIONAL, INC. (US)

CDS

```
700...xref="taxon:35935".
21...908
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD52957.1"
/db_xref="GI:23559678"
/db_xref="R05TREMEL:CAD52957"
/translation="MLKAPDHQVGAHEAGLKGKLPILDDSGRYFKPLOGDNGSBEVA
FRESTSNNILPEHKRFPIYGIKIMKASTGSDHPHWLQDLTAHYNPVSMVDIKI
GRTWAPASEAYIAKCLKQKDESTSIPLGFIAGLVQVDDSGSGYKPHRNVMRTKG
PADRVLLKPFYSSNPASMEMRNTGLGPDCLASFIYGGPNGILQALMKLTKWFEDQT
IYHFVAGSPFLFIYFEKRLVLKGAARSNAEVLIIIDFHVHTDGNGVIDHNFLGLGSLIKFI
SDILSETGDCNCTNGCVET."

```

```

ORIGIN
Alignment Scores:
Pred. No.: 3 02e-47
Score: 742.50
Percent Similarity: 67.83%
Best Local Similarity: 52.10%
Query Match: 48.82%
DB: 6
Length: 1020
Matches: 149
Conservatives: 75
Mismatches: 49
Indels: 13
Gaps: 5

```

US-10-042-894A-8 (1-289) x AX513576 (1-1020)

```
QY 4 LeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyPro 23
   |||
   |||
   |||
Db 24 CTCAGGCCCCAGATCATCAGGTGCTGCATGACATGAAGCTGGGCTGGGAAGCTTGGCCCA 83
   |||
   |||
   |||
QY 24 LeuLeuAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis 43
   |||
   |||
   |||
Db 84 CTCATTGATGATTCAGGCCGGTTCACAAACCACTGCAGGGTGATTAACCGTGGGTCAGAA 143
   |||
   |||
   |||
QY 44 GluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArgLleArgasp 63
   |||
   |||
   |||
Db 144 GAAGTAGCCCTTTATGAATCATCTTCTTAACAATAATATTCCAGAAACACATACGC--- 200
   |||
   |||
   |||
QY 64 ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83
   |||
   |||
   |||
Db 201 AATATCTTCTCTATATATTATGGCACCACAAATCATG-----AAGCATCCCTGGCTCT 254
   |||
   |||
   |||
QY 84 ProHisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaProCysValAla 103
   |||
   |||
   |||
Db 255 GACCATCTCTACATGCTGTTCGAAGATCTTACATCAGCTCATGTCAACCCATCTGTAATG 314
   |||
   |||
   |||
QY 104 AspLleLysLleGlyAlaLleThrTrpProProSerSerProGluProTyrLleAlaLys 123
   |||
   |||
   |||
Db 315 GACATCAAAATCGGCTCCAGACATCGGGCGCCAGAAAGCTTCCGAGCGGTACATTCGAAA 374
   |||
   |||
   |||
QY 124 CysLeuAlaLysAspArgGlyThrThrSerValLeuLeuLeuPheArgValSerGlyVal 143
   |||
   |||
   |||
Db 375 TGTCTTAAAGAGATAGGAAGACCAAGTATTCATTTGGGATTCAGGATTCGGGGCTG 434
   |||
   |||
   |||
QY 144 ArgValValGlyProGluGlyAla---ValTrpArgThrGluArgProGluValLysAla 162
   |||
   |||
   |||
Db 435 CAAGTCTATATCGATGATGGGTGAGGCTTTTATAAGCCTCATAGAAATTTACATGCGTAA 494
   |||
   |||
   |||
QY 163 MetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSerValAlaAspGlu 182
   |||
   |||
   |||
Db 495 ACCGGCCAGCTGATGTAGACTACTCTCTAGGAAATTTGTTCTTCTAACCCGCTGCA 554
   |||
   |||
   |||
QY 183 -----GlyMetAspCysAlaLeuAlaAlaValTyrGly 194
   |||
   |||
   |||
Db 555 GAGATGGAAATCGGCACAGGCCCTAGCCCGGATGTCTTTAGCATCTTTGTTATGT 614
   |||
   |||
   |||
QY 195 GlyLysGlyValLysSerGlnLeuArgGluLeuLysAlaTyrPheGluGlnThr 214
   |||
   |||
   |||
Db 615 GGGCCTAATGGGATATTAGTCAACTGATGGAATTTGAACATGTTTGAAGATCAACA 674
   |||
   |||
   |||
QY 215 LeuPheHisPheTyrSerAlaSerLleLeuLeuGlyTyrAspAlaAlaValAlaAla 234
   |||
   |||
   |||
Db 675 ATTTACCACCTCTAGCTGTGTTCTTTTGTTCATCTTTGAAAGAGAGGTGGTGTAA 734
   |||
   |||
   |||
QY 235 GlyGlyAspGlyGlyValThrValLysLeuValAspPheAlaHisValAlaGluGly 254
   |||
   |||
   |||
Db 735 GGTGCT---CGGTCAACCGCAGAACTCAACTTATGTTGCTCATGTACATGATG 791
   |||
   |||
   |||
QY 255 AspGlyValLleAspHisAsnPheLeuGlyGlyLeuCysSerLeuLleLysPheValSer 274
   |||
   |||
   |||
Db 792 AATGTGTTATGATCACAATTTCTTGGTGGGCTCTGTTCTTTCATGATAAGTTTCATCT 851
   |||
   |||
   |||
QY 275 AspLleValProGluThr 280
   |||
   |||
   |||
Db 852 GACATACCTTTCGGAGACA 869
   |||
   |||
   |||
```

RESULT 11

```
AX513574
LOCUS AX513574
DEFINITION Sequence 11 from Patent WO02059324.
ACCESSION AX513574
VERSION AX513574.1 GI:23559674
KEYWORDS Eucalyptus grandis
SOURCE Eucalyptus grandis
ORGANISM Eucalyptus grandis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
```

rosids; Myrtales; Myrtaceae; Eucalyptus.

REFERENCE

1 Shi, J., Beach, L.R., Wang, H., Rafalski, J.A. and Cahoon, R.E.
Novel inositol polyphosphate kinase genes and uses thereof
Patent: WO 02059324-A 11 01-AUG-2002;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES

source

1. 1195
/organism="Eucalyptus grandis"
/mol_type="unassigned DNA"
/db_xref="taxon:71139"
116. 1048
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD52956.1"
/db_xref="GI:23559675"

CDS

116. 1048
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD52956.1"
/db_xref="GI:23559675"
/translation="MLKVPDQVAGHGGKGLPLVDSDGRFYKPLQSDHRGDTVEA
FYEFYSNTEIPGHIRKEFPFHGKTKEASDGPQPHLVLIEDVSGRTNPDLMDIK
TGSRTWYPEASEEYIOKLEKDRNSTSVSLGRISGLRVYONSEAGFWQPKKVVYF
NADGVSALRKFEVSNLSLGNVDPDCHYASKYCHRGGLAQLQLKEWFEVQNYH
FYSCSLIYDRESALDGCARPKVKLVDFAHVMDGHGVIDHNFLGLGSLVIRKIDIA
DEDNKACKEVNLGLKENGFKYSSTPELDHEAC"

ORIGIN

Alignment Scores:

Pred. No.: 1,18e-45 Length: 1195
Score: 722.50 Matches: 149
Percent Similarity: 66.31% Conservative: 38
Best Local Similarity: 52.84% Mismatches: 84
Query Match: 47.50% Indels: 11
DB: 6 Gaps: 6

US-10-042-894A-8 (1-289) x AX513574 (1-1195)

```
QY 4 LeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyPro 23
   |||
   |||
   |||
Db 119 CTCAGGTCCTGGATCATCAAGTCGCGGTCCACGGGGAGAGCGGGGAAAGCTGGGCCA 178
   |||
   |||
   |||
QY 24 LeuLeuAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis 43
   |||
   |||
   |||
Db 179 CTGGTGATGATTCGGGCGCTCTTATAAGCTCTCCAGAGCGATCATCGCGAGACAG 238
   |||
   |||
   |||
QY 44 GluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArgLleArgasp 63
   |||
   |||
   |||
Db 239 GAAGTGGCCTTTACGAGTCATTCATTCATACCCGAGATCCCGAGTCACATTCGC--- 295
   |||
   |||
   |||
QY 64 ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83
   |||
   |||
   |||
Db 296 AAATCTTCTCTGCTTTCACGGAACCTAAGACTATT-----GAGGCGCTCTGATGATCG 349
   |||
   |||
   |||
QY 84 ---ProHisProHisLeuValLeuAspPheLeuAlaGlyPheGlnAlaProCysVal 102
   |||
   |||
   |||
Db 350 GGTCTCAACCTCACTGCTGTTCTGGAGGATCTGCTCGGGTCGCGAGGCCATCTCTC 409
   |||
   |||
   |||
QY 103 AlaAspLleLysLleGlyValAlaLleThrTrpProProSerSerProGluProTyrLleAla 122
   |||
   |||
   |||
Db 410 ATGACATCAAGACTGGATCCAGAACATGGTATCCGAGGCGCTCTGAGAGGTACATCA 469
   |||
   |||
   |||
QY 123 LysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGly 142
   |||
   |||
   |||
Db 470 AAGTGTCTTAGAGAAAGATCGAAATAGCACAACGCTTTCATTTGGGTTTTAGGATTTCTGG 529
   |||
   |||
   |||
QY 143 ValArgValVal---GlyProGluGlyAlaValTrpArgThrGluArgProGluValLys 161
   |||
   |||
   |||
Db 530 CTAAAGGGTATATCAAAATAGCAAGCTGGATTTTGGCAACCTGAGAAAGAGGTTGTTAT 589
   |||
   |||
   |||
QY 162 AlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSerValAlaAsp 181
   |||
   |||
   |||
Db 590 AGCTTTAATGCGGCGGTGTGAGTCCGCTCTGAGGAAGTTGTTCTTCTCAACTGTCT 649
   |||
   |||
   |||
QY 182 GluGly-----MetAspCysAlaLeuAlaAlaValTyrGlyLysGly 197
   |||
   |||
   |||
```



```
PSLAELTKLHLKLVIGQCNIIGIYGLANLGSLSHLEISGCPCLTVVBSLRLNLIGT
LELSGYTITNLPBSLTYTKLSKLVSDSOLPDLTNLKNLRCIKICSDNFIEITGL
HTLSLELRVWGSIRKLDLTGLVKLELLAFDSTQITETIRGGLSRLRHSRC
QSIRELPNLSGLKLSYILLEKCRHLKEVIRLEELKRLDFNTHRG"
17600..18457
/function="unknown"
/note="ORF
related to human inositol hexakisphosphate kinase 2
protein"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAC35324.1"
/db_xref="GI:13509210"
/db_xref="GOA:Q9AR40"
/db_xref="SPTREMBL:Q9AR40"
/translation="MLKVPKQVAGHQDINGLGLPLVDSDGRFYKPLQCDREGATEMA
FYTSITNKTIPOKVRREFQAFQGTQSEASDGLPHLLIEDTLRLHLVMDIK
IGSRVYPEASQAYIEKCLKXDVESNPLLFRI.SGLQVNGKSESEVWLKPRKLL
LNLNADEVLRVFNKSVNLSDDPCDFAAVSVSGNGLLAQLLEKAWPEDQITVH
FNSGSLVLYEKEKTKMVNGGSESLGGAACKLIDFAHVTGNGVIDHFNFLGLSLIK
FISENQIML"
join(20265..20815,21104..22202,22415..22690,22818..24218)
/gene="Ngc-B"
Join(20265..20815,21104..22202,22415..22690,22818..24218)
/gene="Ngc-B"
/note="N rust resistance gene homolog: TIR-NBS-LRR
protein"
/codon_start=1
/product="Ngc-B protein"
/protein_id="CAC35325.1"
/db_xref="GI:13509211"
/db_xref="GOA:Q9AR40"
/db_xref="SPTREMBL:Q9AR40"
/translation="MRSDDSSICGSFLRSVSDPTLPLPSGEYEVFLSPRGPVDRQ
TFADHLXAWLVKSLIRTPDEEGKQGETIGSSLIQATISKIYIPILITQNYASSKW
LELAKWVDKWNKGKGGQKHIIIPVFYFMDPRVDRHPDSGPYKEAFQHNKHDPET
ILEWKEALQDVYKMGKWHINELTQCGAVVDKIFTTIRPHLRNVTLATDELVGDDSV
EVMELMNLDSHSTSEIRIIGYMGGLGKNTLAKAVENKVSQPERCCFLDNIREFLR
NDGVVALQNKVSIIDILRKSQAKNASDGVLIIBERVRRHKIFVVLDDIBSEPHFDEI
FGKLGDSFTRSLITRDARTLELNECKMFGLEMSHDSHLSQLFKHAPGVDPPE
DYASLCREFIVASGLPALKVIGSLIFKSDRFWEKLELKAIPSAKVQOERLKVSY
LMTNHEKQIFLDIACLVGAKYKVPYMSDCLYIPASTLRTLVQSLVLRMDNKKF
WMHDHIDRLGRTIYRENSQNPYKRSIWSNNDAILKRNREGNDCVEALRVDMMKGG
YALTNKQFQKSRFLVLENDLSGNFKNLPNLRVVRVYRGPSPSGNLNPKLVIL
ELDCVYTHSGWNEIKAAGLKVNLITSGILEKVPDLSTCGELCLPHKQWNR
GELDITFKDLKVLIDNQETITLKGVESLQNLQDLVGRSGLIEVPAGISKUSLE
FLDLSVAKDEVEMLPGLKLVISFSLSALPSSLIKLDICDSRNLRQPNLASVTN
LTRHLKVEHGIPEGLGKLVLESLSICNAPNDNDLGDLENLVLLKELALERPCLIG
KLPSAELTKLHVIRWCDVLGIYGLGNLGDLSHLDISWCPRLTVMOLLHSLIKL
GTIVSSGFELTNILPLSLIYTKLRTLEVRSSQLPDLTNLKNLRDLITITGCRELIETL
GLHTLESLELSMERCPSPVKLDIAGLIKLTIIHICTOLTETIRGGLGLESLOLPM
SGCGSIKELNLSGLKLVKYPKLECRQLKEVNGLEBLEWLDNFNDRRLKULKLV
SRKQQLVTSARRALFINN"
ORIGIN
Alignment Scores:
Pred. No.: 8,32e-44 Length: 25054
Score: 718.50 Matches: 149
Percent Similarity: 66.78% Conservative: 44
Best local Similarity: 51.56% Mismatches: 83
Query Match: 47.24% Indels: 13
DB: 8 Gaps: 7
US-10-042-894A-8 (1-289) x LUS310150 (1-25054)
Qy 7 ProGluHisGlnValAlaGlyHisArgAlaSerLysLeuGlyProLeuLeuAsp 26
Db 9625 CCGGAACATCAGGTGCTGGACATCAAGCCATCAACGGCCCTCTCGTCTCTTGTAGAC 9684
Qy 27 GlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHisGluValAla 46
Db 9685 GATTGAGGCGGATTCTACAGCCACTTCAGGCGCATGAACGTGGGCAACCGAGATGCCA 9744
```

```
Qy 47 PheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArgIleArgAspThrPhePhe 66
Db 9745 TTCACACATCTTTTCGTCTCAAAAACGATTCACAAACACGTCGG---AGGTTTTC 9801
Qy 67 ProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGluPro---His 85
Db 9802 CCGGCCTTCATGCGACTCAAAAGTATA-----GAGGCTTCTGATGGATCTGGCCTCCAC 9855
Qy 86 ProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaProCysValAlaAspIle 105
Db 9856 CCGCACCTCATCTTAGAAGATCTCACGTCGATCGCTGGTTGCATCCATGTGTATGACATC 9915
Qy 106 LysIleGlyAlaIleThrTyrProSerSerProGluProTyrIleAlaLysCysLeu 125
Db 9916 AAGATTGGTCCAGACACATGTCACCCGAGGCTTCTCAAGCTTACATCGAAGATGCTG 9975
Qy 126 AlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyValArgVal 145
Db 9976 AAGAAGGATGTGGAATCAAGCAGTCCCTTCTCGGGTTTAGGATATCCGGATTGCGAGTT 10035
Qy 146 ValGlyProGlu-----GlyAlaValTrpArgThrGluArgProGluValLys 161
Db 10036 TACGGTAACGACAAAGAAATCATCTGAGGTGCTCTTGAAGCCTGAGAGGAGCTTCTCCAG 10095
Qy 162 AlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSer---ValAla 180
Db 10096 AATCTTACTGCTGATGAAGTTAGCTGCTGTGCTTAAACGGTTGTTCTTCCACCCGAG 10155
Qy 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyValLeu 200
Db 10156 TCGGACCAACCGGATTGTTGCTTTGACGCGTCTGTTACGGTGGTTCGAATGGGATCTTG 10215
Qy 201 SerGlnLeuArgGluLeuLysAlaTrpPheGluGlnThrLeuPheHisPheTyrSer 220
Db 10216 GCTCAGTTGCTGGAGCTGAAGCGGTGTTGAGGATCAGACCATCTATCATTTCAACTCT 10275
Qy 221 AlaSerIleLeuLeuGlyTyr-----AspAlaAlaAlaValAlaAlaGlyValAspGly 238
Db 10276 TGTCTCTGTTCTTATGCTGTATGAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 10335
Qy 239 -----GlyGlyValThrValLysLeuValAspPheAlaHisAlaGluGlyAspGly 256
Db 10336 AGTTTGGGTTGCTGTCGGGTGAAGCTCATCGATTTTCTCATGTGACGGAGGAAATGGC 10395
Qy 257 ValIleAspHisAnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIle 276
Db 10396 GTTATCATGATCACAACTTCTTGGAGGCTCTGCTCATTTGATAAAAGTTTCATTTCGAGATC 10455
Qy 277 ValProGluThrProHisThrGlnPro 285
Db 10456 CTCACCGTCGGATGAGAAATCCCG 10482
RESULT 13
ATH243592 903 bp DNA linear PLN 01-JUL-2001
LOCUS Arabidopsis thaliana gene for inositol-(1,4,5) trisphosphate
DEFINITION 3-kinase.
ACCESSION AJ243592
VERSION AJ243592.1 GI:14588984
KEYWORDS inositol-(1,4,5) trisphosphate 3-kinase.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1
Xia,H.J., Brearley,C. and Mueller-Roeber,B.
Identification of inositol-(1,4,5) trisphosphate 3-kinase from
Arabidopsis thaliana
Unpublished
JOURNAL 2 (bases 1 to 903)
REFERENCE 2
AUTHORS Xia,H.J.
TITLE Direct Submission
```

JOURNAL Submitted (09-JUL-1999) Xia H.J., AG Mueller-Roeber,
Max-Planck-Institute of Molecular Plant Physiology,
Karl-Liebknecht-Strasse 25, Haus 20, 14476 Golm, GERMANY

FEATURES Location/Qualifiers

source 1..903

/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Col"
/db_xref="taxon:3702"

CDS 1..903

/function="phosphorylation of the second messenger
inositol-(1,4,5) trisphosphate"
/codon_start=1
/product="inositol-(1,4,5) trisphosphate 3-kinase"
/protein_id="CAC43070.1"
/db_xref="GI:14588985"
/db_xref="GOA:Q9FLT2"
/db_xref="SPTREMBL:Q9FLT2"
/translation="MLKVPBPVHVAGHISADGKGLPIVDQGRFFKPLQDGRGEHEAK
FYSEFTSNMKVPDHIIRYFVYHTQGLVEASDGSGLPHLDVLDVVGVYANPSVMVDV
IGRTWYLPDYSEYFKKCIKKDKTTTSLVSPRVSGFKIPFDHSESFWRAEKKLVLYG
NADGALPDKRFVSSPNKDSNLTNCPAFSEVYGCNGILAQLLELKQWFEQTQTLVY
FNCSILMIYENSIILMOGDDAPAPRAQVKLVDPFHVLDGNGVIDHNFLGGLCSFIK
FIKDIQLQSEKHETTSLENGR"

ORIGIN

Alignment Scores:

Pred. No.:	3,41e-45	Length:	903
Score:	714.50	Matches:	150
Percent Similarity:	66.55%	Conservative:	39
Best Local Similarity:	52.82%	Mismatches:	82
Query Match:	46.98%	Indels:	13
DB:	8	Gaps:	8

US-10-042-894A-8 (1-289) x ATH243592 (1-903)

QY	4	LeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLysLeuGlyPro	23
Db	4	CTCAAGTCCCTCAACACCAAGTTGCTGGTCACATTGCTAGTCATGGGAAGCTCGGTCCA	63
QY	24	LeuLeuAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis	43
Db	64	CTCGTAGATACCAAGCCGGTCTTCAAGCCACTTCAGGAGATTCGTGGCGCAACAC	123
QY	44	GluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArgIleArgAsp	63
Db	124	GAGGCTAAGTTCTATGAGCTTTTCATCGAACATGAAGTTCACATCATCCAT---	180
QY	64	ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu	83
Db	181	AGATACTTCCGGTGTATCAGCGCACTCAGCTAGTT-----GAAGCATCTGATGATCT	234
QY	84	ProHis--ProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaProCysVal	102
Db	235	GGCAAGTTCCTCATCTGTTCTTGATGATGTTGTTTCAGGATACGGAACCCGTCGGTA	294
QY	103	AlaAspIleLysIleGlyAlaIleThrTrpProProSerSerProGluProTyrIleAla	122
Db	295	ATGGATGTTAAGATTGGATCTAGACATGGTACCCTCGGATCTATCAGAAGAACTACTCAAG	354
QY	123	LysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGly	142
Db	355	AAATGTTATTAAAGAAATAGACAGACACCACCGTTTCGTTGGGGTTTCAGGGTTTCAGGT	414
QY	143	ValArgValValGly--ProGluGlyAlaValTrpArgThrGluArgProGluValLys	161
Db	415	TTTAGATTTTTCATCACCAGGAATCAAGTTTTCAGGAGCTGAGAAGAAGCTTGTCTCT	474
QY	162	AlaMetAspThrAlaGlyValArgGlyValLeuArgArgTyrValSer-----SerVal	179
Db	475	GGGTATATGCATAGTGTCTAGATTGGCTCTGAGGAAGTTTGTGATCGAACTCTCC	534
QY	180	AlaAspGluGlyMet-----AspCysAlaLeuAlaAlaValTyrGlyGlyLysGly	197

Theologis,A., and Ecker,J.R.

Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES

source

Location/Qualifiers

1..903

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="5"

/clone="U13510"

/ecotype="Columbia"

/note="This clone is in pUNI 51"

1..903

/note="unknown protein"

/codon_start=1

/product="AT5G61760/mac9_50"

/protein_id="AAL62012.1"

/db_xref="GI:18252263"

FYESTSNMKVDPHRIYFPVYHGTQVLEASDGSGLPHLVLDVVGVPVNDVK
IGRTWYPOVSSEYFKCIKKDQRTTTSVLSGRVSGFKIFDQESSFWRAEKLVLGY
NADGARLARKEVSSNSPADSNLTPNCAPASEVYGGCNGILAQLELQWFTQTLVH
FNSCSILMIYENESILMQGGDAPAPRAQVKLVDFAHVLDGNGVIDHNFGLGCSFIK
FIKDILQSVKXHDFTDTSLENGR"

ORIGIN

Alignment Scores:

Pred. No.: 3 41e-45 Length: 903
Score: 714.50 Matches: 150
Percent Similarity: 66.55% Conservative: 39
Best Local Similarity: 52.82% Mismatches: 82
Query Match: 46.98% Indels: 13
DB: 8 Gaps: 8

US-10-042-894A-8 (1-289) x AY072621 (1-903)

QY 4 LeuHisProGluHisGlnValAlaGlyHisArgAlaSerLeuGlyPro 23
DB 4 CTCAGGTCCTCGAACACCAAGTGTCTGTACATGCTAGTGGGAGCTCGTCA 63
QY 24 LeuLeuAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis 43
DB 64 CTCGTAGATGACCAAGCGCGTCTCTCAAGCCACTTCAGGAGATCTCGTGGCGAACAC 123
QY 44 GluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArgIleArgasp 63
DB 124 GAGGTAAGTCTTATGAGTCTTTCATCGAACATGAGGATTCAGATCATCATCAT 180
QY 64 ThrPheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGlu 83
DB 181 AGATACCTCCCGGTGTATCAGCGCACTCAGTAGT-----GAGCATCTCATGATCT 234
QY 84 ProHis---ProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaProCysVal 102
DB 235 GCAAGCTCTCTCATCTCTGTTCTTCTGATGATGTTGTTTTCAGGATGACCAACCCGCGTA 294
QY 103 AlaAspIleLeuValLeuAlaIleThrTrpProSerSerProGluProTyrIleAla 122
DB 295 ATGGATGTTAAGATTGGATCTAGGACATGTTACCGGATGATACAGAAATCTTCAG 354
QY 123 LysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGly 142
DB 355 AAATGATTAAAGAAAGATAGACAGACCACCGCTTCGTGGGTTTCAGGTTTCAGGT 414
QY 143 ValArgValValGly---ProGluGlyAlaValTrpArgThrGluArgProGluValLys 161
DB 415 TTTAAGATTTTTCATCACCAGAAATCAAGTTTTCGAGAGCTGAGAAAGCTTCTTCT 474
QY 162 AlaMetAspThrAlaGlyValArgValLeuArgValLeuArgTyrValSer-----SerVal 179
DB 475 GGGTATATGCGAGTGGTCTAGATGGCTCTGAGGAAGTTTGTGTCTCGAATCTCC 534

QY 180 AlaAspGluGlyMet-----AspCysAlaLeuAlaAlaValTyrGlyGlyLysGly 197
DB 535 GCTGACTCTTAACCTTGACACCAACTGTCTTTTGCATCAGAGGTTTATGCGGTGTGAAC 594
QY 198 GlyValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHis 217
DB 595 GGGATCTTAGCGCAGTTGTTGGAGCTTAAAGATTGGTTCGAAACCCCAAGCCTTACCAT 654
QY 218 PheTyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaVal---AlaAlaGlyGly 236
DB 655 TTTCAATCTCTGCTCGATTTCTGATTTACGAGATGAATCAATCTTGTATGCAAGGAGGA 714
QY 237 AspGlyGly-----GlyValThrValLysLeuValAspPheAlaHisValAlaGlu 253
DB 715 GATGATGACCGGACCGGACCAAGTGAAGCTGGTGAATTCGCGATGTTCTTGAT 774
QY 254 GlyAspGlyValIleAspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLysPheVal 273
DB 775 GGAACCGGTGTCATCGACCATTAATTTCTTGGGTGGACTCTGCTCTTCATAAGTTCATC 834
QY 274 SerAspIleVal 277
DB 835 AAAGATATTTCTT 846

RESULT 15

AY147936

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AY147936 903 bp mRNA linear PLN 10-NOV-2002
Arabidopsis thaliana inositol polyphosphate 6-/3-/5-kinase 2b
(IPK2b) mRNA, complete cds.
AY147936.1 GI:24850180
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
Stevenson-Paulik,J., Odom,A.R. and York,J.D.
Molecular and Biochemical Characterization of Two Plant Inositol
Polyphosphate 6-/3-/5-Kinases
J. Biol. Chem. 277 (45), 42711-42718 (2002)
12226109
2 (bases 1 to 903)
Stevenson-Paulik,J.M. and York,J.D.
Direct Submission
Submitted (04-SEP-2002) Pharmacology and Cancer Biology, Howard
Hughes Medical Institute, Duke University Medical Center, LSRC
Building C266, Research Drive, Durham, NC 27710, USA
Location/Qualifiers
1..903
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="5"
/ecotype="Columbia"
1..903
/gene="IPK2b"
1..903
/note="inositol phosphate kinase"
/codon_start=1
/product="inositol polyphosphate 6-/3-/5-kinase 2b"
/protein_id="AAN63056.1"
/db_xref="GI:24850181"

FEATURES

source

gene

CDS

ORIGIN

Job time : 4547 secs

Alignment Scores:

Pred. No.:	3,41e-45	Length:	903
Score:	714.50	Matches:	150
Percent Similarity:	66.55%	Conservative:	39
Best Local Similarity:	52.82%	Mismatches:	82
Query Match:	46.98%	Indels:	13
DB:	8	Gaps:	8

US-10-042-894A-8 (1-289) x AY147936 (1-903)

QY	4	LeuHisProGlnHisGlnValAlaGlyHisArgAlaSerLysLeuGlyPro	23
DB	4	CTCAAGTCCTCGAACCAACCAAGTTGCTGTCACATTGCTAGTATGGGAAGCTCGGTCCA	63
QY	24	LeuLeuAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArgGlyGluHis	43
DB	64	CTCGTAGATGACCAAGGCCGGTCTTCAAGCCACTTCAGGGAGATTCTCGTGGCGAACAC	123
QY	44	GluValAlaPheTyrGlnAlaPheSerAlaHisAlaValProAlaArgIleArgAsp	63
DB	124	GAGGCTAAGTCTCATGAGTCTTTCACATCGACATGAAGGTTCCAGATCACATCCAT---	180
QY	64	ThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGlnAlaGlnProGlyGlu	83
DB	181	AGATACCTCCCGTGTATACCGGCATCTAGCTAGTT-----GAAGCATCTGATGGATCT	234
QY	84	ProHis---ProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaProCysVal	102
DB	235	GGCAAGCTTCCTCATCTTGTTCTTGATGATGTTGTTGAGGTACGCAACCCGTCGGTA	294
QY	103	AlaAspIleLysIleGlyAlaIleThrTrpProSerSerProGluProTyrIleAla	122
DB	295	ATGGATGTTAAGATTGGATCTAGGACATGTCACCGGATGATCAGAGATACTTCAAG	354
QY	123	LysCysLeuAlaLysAspArgGlyThrSerValLeuLeuGlyPheArgValSerGly	142
DB	355	AAATGATTAAAGAAAGACAGACACACCGCTTTCGTTGGGGTTTCAGGGTTTCAGGT	414
QY	143	ValArgValValGly---ProGluGlyAlaValTyrArgThrGluArgProGluValLys	161
DB	415	TTTAAGATTTTCATCACCAGATCAAGTTTGGAGGCTGAGGAAGCTTGTCTT	474
QY	162	AlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSer-----SerVal	179
DB	475	GGGTATAATGCAGATGGTGTAGATTGGCTCTGAGGAAGTTTGTGTCATCGAACTCTCCC	534
QY	180	AlaAspGluGlyMet-----AspCysAlaLeuAlaAlaValTyrGlyGlyLysGly	197
DB	535	GCTGACTCTAAGTACACCAACTGCTTTTGCATCAGAGGTTTATGGCGTTGTAAC	594
QY	198	GlyValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHis	217
DB	595	GGGATCTTAGCGCAGTTGTTGGAGCTTAAAGATTGTTGMAACCAACGCTTTACCAT	654
QY	218	PheTyrSerAlaSerIleLeuLeuGlyTyrAspAlaAlaVal---AlaAlaGlyGly	236
DB	655	TTCAATTCTGCTCGATTCTGATGATTACGAGATGAATCAATCTTGATGCAAGGAGGA	714
QY	237	AspGlyGly-----GlyValThrValLysLeuValAspPheAlaHisValAlaGlu	253
DB	715	GATGATGCCCGGCACACCGGCACAAAGTGAAGTGGTGAATTCGCGCATGTTCTTGAT	774
QY	254	GlyAspGlyValIleAspHisAnPheLeuGlyGlyLeuCysSerLeuIleLysPheVal	273
DB	775	GGAAACGGTGTATCGACCATATTTCTGGGTGGACTCTGCTCTTTTCATAAAGTTTCATC	834
QY	274	SerAspIleVal	277
DB	835	AAAGATATTCTT	846

Search completed: March 27, 2004, 07:35:02

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2004, 20:51:49 ; Search time 18 Seconds

(without alignment)

836.015 Million cell updates/sec

Title: US-10-042-894A-8

Perfect score: 1521

Sequence: 1 MSDLHPPEHQVAGHRASAK.....IKFVSIVPETHPTQLPLGS 289

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	145	9.5	461	1	IP3K HUMAN
2	131	8.6	459	1	IP3K RAT
3	123	8.1	355	1	ARG3 YEAST
4	114.5	7.5	872	1	IP3L RAT
5	109.5	7.2	946	1	IP3L HUMAN
6	97.5	6.4	362	1	ADD_MYCLE
7	93.5	6.1	372	1	CARA_SULTO
8	89.5	5.9	331	1	NAGZ_XANCP
9	89	5.9	380	1	EGSA_THEAC
10	88.5	5.8	352	1	NRJ1 STRPE
11	88.5	5.8	370	1	VRXJ_AGRF5
12	88.5	5.8	377	1	VRXJ_AGRF5
13	87.5	5.8	510	1	ER53 CERAE
14	87.5	5.8	550	1	CHIT NPVOF
15	86.5	5.7	373	1	CSZ MYCTU
16	86.5	5.7	377	1	VRXJ_AGRF9
17	86.5	5.7	454	1	UCR2_NEUCR
18	86.5	5.7	1032	1	YGFK_ECOLL
19	86	5.7	364	1	X80 MYCTU
20	85.5	5.6	334	1	NAGZ_XANAC
21	84.5	5.6	447	1	DCDA MYCTU
22	84	5.5	316	1	PECT_ERWCH
23	84	5.5	411	1	DRE3_ARATH
24	83.5	5.5	304	1	POQB_PSEAE
25	83.5	5.5	1201	1	MED_WYXAA
26	83	5.5	131	1	PRO3_LILLO
27	83	5.5	492	1	DP22 HUMAN
28	82.5	5.4	485	1	ARLY_HALN1
29	82	5.4	131	1	PRO2_ARATH
30	82	5.4	310	1	MFAA_ERUSU
31	82	5.4	541	1	G6P1_DEIRA
32	82	5.4	742	1	UL47_HSVBP
33	81.5	5.4	529	1	IMDH_MYCTU

34 81.5 5.4 884 1 RPOB_NPVOP O12934 oryza pseu
35 81 5.3 411 1 DHE1_ARATH Q43314 arabidopsis
36 80.5 5.3 298 1 KHSE_ANASP Q8VZV9 anabaena sp
37 80.5 5.3 352 1 ADH_THERB P14941 thermoanaer
38 80.5 5.3 354 1 TRD1_STRGO O68608 streptomyce
39 80.5 5.3 391 1 Y532_METJA Q57952 methanococ
40 80.5 5.3 1364 1 PUR2_DROPS P16340 d trifuncti
41 80 5.3 294 1 PANE_ARCFU O28578 archaeoglob
42 80 5.3 311 1 MIAA_MYCLE P46811 mycobacteri
43 80 5.3 411 1 DHE2_ARATH Q38946 arabidopsis
44 80 5.3 411 1 DHE3_VITVI P52596 vitis vinif
45 80 5.3 421 1 FAFB_RHIME P56902 rhizobium m

ALIGNMENTS

RESULT 1
IP3K HUMAN
ID IP3K HUMAN STANDARD; PRT; 461 AA.
AC P23677;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Inositol-trisphosphate 3-kinase A (EC 2.7.1.127) (Inositol 1,4,5-
trisphosphate 3-kinase) (IP3K) (IP3 3-kinase).
GN ITPKA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=91128380; PubMed=1847047;
RA Takazawa K., Perret J., Dumont J.E., Erneux C.;
RT "Molecular cloning and expression of a human brain inositol 1,4,5-
trisphosphate 3-kinase.";
RL Biochem. Biophys. Res. Commun. 174:529-535(1991).
[2]
RN SEQUENCE FROM N.A.
RP TISSUE=Brain;
RX MEDLINE=91088302; PubMed=2175886;
RA Takazawa K., Perret J., Dumont J.E., Erneux C.;
RT "Human brain inositol 1,4,5-trisphosphate 3-kinase cDNA sequence.";
RL Nucleic Acids Res. 18:7141-7141(1990).
CC -!- CATALYTIC ACTIVITY: ATP + 1D-myo-inositol 1,4,5-trisphosphate =
ADP + 1D-myo-inositol 1,3,4,5-tetrakisphosphate.
CC -!- ENZYME REGULATION: IP3K is activated by calmodulin.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X54938; CAA38700.1; .
DR PIR; JN0129; JN0129.
DR Genew; HGNC:6178; ITPKA.
DR MIM; 147521; .
DR GO; GO:0008440; P:inositol-trisphosphate 3-kinase activity; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR005522; IPK.
DR Pfam; PF03770; IPK; 1.
DR TRANSFERASE; Kinase; Calmodulin-binding.
FT DOMAIN 287 295 CALMODULIN-BINDING (BY SIMILARITY).
SQ SEQUENCE 461 AA; 51008 MW; 18CA214A091F5B19 CRC64;
Query Match 9.5%; Score 145; DB 1; Length 461;
Best Local Similarity 23.5%; Pred. No. 0.0001;
Matches 76; Conservative 40; Mismatches 97; Indels 110; Gaps 16;

QY 10 QVAGHRASAKGLPLDGSGLFYKPLQAGDRGEHEVAFYAFSAHAAVAPARIRDTFFPRF 69
 Db 190 QLAGHTGSFKAAGT---SGLLKRCSEPER--YCLARLMADALRGCVPA-----F 234
 QY 70 HGTRELLPTEAQCPGPHPLVLDLLAGFQAPCVADIKIGAITW----- 112
 Db 235 HG-----VVERDGB--SYLQQLDLDGDFGCVLDCKMGVRYLYEELTKARERPKLRKD 287
 QY 113 -----PPSSP--EPYIAKCLAKDRG-----TTSVLLGFRVSGVRVV--GPEGAVVR 154
 Db 288 MYKMLAVDPEAPTEEHEAQRVTKRYMQWREGISSSTTLGFRIGIKKADGSCSTDFK 347
 QY 155 TERPEVKAMD-----TAGVRVLRVRYSSVADEGMDCALAAAVYGGKGVLSQRLKAW 209
 Db 348 TTRSREQVTRVFEEFVQGDDEVLRV-----LNRLQIQRDT 383
 QY 210 FEEQTLFHFY---SASILLGVDAAAVAGDGGGVTVKLVDAEHAFAE--GDG-VIDH----- 260
 Db 384 LEVSEFFRHEVIGSSLLFVHDCHHRAG-----VWLIDFGKTTPLPDGQILDHRRPW 435
 QY 261 -----NFLGGLCSLIKVFSDI 276
 Db 436 EEGNREDGYLLGLDNLIGILASL 458
 RESULT 2
 IP3K_RAT
 ID IP3K_RAT STANDARD; PRT; 459 AA.
 AC P17105;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Inositol-trisphosphate 3-kinase A (EC 2.7.1.127) (Inositol 1,4,5-
 DE trisphosphate 3-kinase) (IP3K) (IP3 3-kinase).
 GN ITPKA.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=101116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91090700; PubMed=2176078;
 RA Takazawa K., Vandekerckhove J., Dumont J.B., Erneux C.;
 RT "Cloning and expression in *Escherichia coli* of a rat brain cDNA
 RT encoding a Ca²⁺/calmodulin-sensitive inositol 1,4,5-trisphosphate
 RT 3-kinase.";
 RL Biochem. J. 272:1107-1112 (1990).
 RN [2]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=90208336; PubMed=2157285;
 RA Choi K.Y., Kim H.K., Lee S.Y., Moon K.H., Sim S.S., Kim J.W.,
 RA Chung H.K., Rhee S.G.;
 RT "Molecular cloning and expression of a complementary DNA for inositol
 RT 1,4,5-trisphosphate 3-kinase.";
 RL Science 248:64-66 (1990).
 CC -!- CATALYTIC ACTIVITY: ATP + 1D-myo-inositol 1,4,5-trisphosphate =
 CC ADP + 1D-myo-inositol 1,3,4,5-tetrakisphosphate.
 CC -!- ENZYME REGULATION: IP3K is activated by calmodulin.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X56917; CAA40248.1; -
 DR EMBL; M29787; AAA41457.1; ALT_INIT.
 DR PIR; S13064; S13064.
 DR InterPro; IPR005522; IPK.

DR Pfam; PF03770; IPK; 1.
 KW Transferase; Kinase; Calmodulin-binding.
 FT DOMAIN 285 293 CALMODULIN-BINDING.
 SQ SEQUENCE 459 AA; 50870 MW; 8093DADC0FAD290 CRC64;
 Query Match 8.6%; Score 131; DB 1; Length 459;
 Best Local Similarity 22.6%; Pred. No. 0.0015;
 Matches 73; Conservative 40; Mismatches 100; Indels 110; Gaps 16;
 QY 10 QVAGHRASAKGLPLDGSGLFYKPLQAGDRGEHEVAFYAFSAHAAVAPARIRDTFFPRF 69
 Db 188 QLAGHTGSFKAAGT---SGLI---LKRSEPEHYCL-----VRLMADVLRGCVPAF 232
 QY 70 HGTRELLPTEAQCPGPHPLVLDLLAGFQAPCVADIKIGAITW----- 112
 Db 233 HG-----VVERDGB--SYLQQLDLDGDFGCVLDCKMGVRYLYEELTKARERPKLRKD 285
 QY 113 -----PPSSP--EPYIAKCLAKDRG-----TTSVLLGFRVSGVRVV--GPEGAVVR 154
 Db 286 MYKMLAVDPEAPTEEHEAQRVTKRYMQWREGISSSTTLGFRIGIKKADGSCSTDFK 345
 QY 155 TERPEVKAMD-----AGVRVLRVRYSSVADEGMDCALAAAVYGGKGVLSQRLKAW 209
 Db 346 TTRSREQVTRVFEEFVQGDDEVLRV-----LNRLQIQRDT 381
 QY 210 FEEQTLFHFY---SASILLGVDAAAVAGDGGGVTVKLVDAEHAFAE--GDG-VIDH----- 260
 Db 382 LEISDFFRHEVIGSSLLFVHDCHHRAG-----VWLIDFGKTTPLPDGQILDHRRPW 433
 QY 261 -----NFLGGLCSLIKVFSDI 276
 Db 434 EEGNREDGYLLGLDNLIGILANL 456
 RESULT 3
 ARG3_YEAST
 ID ARG3_YEAST STANDARD; PRT; 355 AA.
 AC P07250;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Arginine metabolism regulation protein III.
 GN ARG3 OR ARG82 OR YDR173C OR YD9395.06C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87257286; PubMed=3239899;
 RA Dubois E., Bercy J., Messenguy F.;
 RT "Characterization of two genes, ARGRI and ARGRIII required for
 RT specific regulation of arginine metabolism in yeast.";
 RL Mol. Gen. Genet. 207:142-148 (1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RA Murphy L., Harris D.E., Barrell B.G., Rajandream M.A.;
 RL Submitted (NCV-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: It is proposed that the ARGRI and ARGRIII proteins are
 CC more likely to be involved in the control of ARGRIII activity.
 CC -!- MISCELLANEOUS: The ARGRI, ARGRII and ARGRIII regulatory proteins
 CC are necessary for the repression of six anabolic genes and the
 CC induction of two catabolic genes by arginine.
 CC -!- MISCELLANEOUS: The expression of this protein is not effected by
 CC the presence of arginine.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X56917; CAA40248.1; -
 DR EMBL; M29787; AAA41457.1; ALT_INIT.
 DR PIR; S13064; S13064.
 DR InterPro; IPR005522; IPK.

RA Bertsch U., Suesse S., Frerik S., Fanick W.;
RT "Cloning of the complete protein coding regions for inositol 1,4,5-
RL triphosphate 3-kinase B-isoforms from rat and human."
RN Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RP [3]
RA SEQUENCE OF 257-946 FROM N.A.
RL Donnelly S.;
RN Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RP [4]
RA SEQUENCE OF 442-946 FROM N.A.
RL TISSUE=Brain;
RX MEDLINE=91378954; PubMed=1654894;
RA Takazawa K., Perret J., Dumont J.E., Erneux C.;
RT "Molecular cloning and expression of a new putative inositol 1,4,5-
RL triphosphate 3-kinase isoenzyme."
RL Biochem. J. 278:883-886(1991).
CC -|- CATALYTIC ACTIVITY: ATP + 1D-myo-inositol 1,4,5-trisphosphate =
CC ADP + 1D-myo-inositol 1,3,4,5-tetrakisphosphate.
CC -|- ENZYME REGULATION: IP3K is activated by calmodulin. Form B
CC is much more sensitive to calcium/calmodulin than form A.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; Y18024; CAB65055.3; -;
DR EMBL; AJ242780; CAC40650.1; -;
DR EMBL; AL365444; CAD20257.1; -;
DR EMBL; X57206; CAA40491.1; ALT_INIT.
DR Genew; HGNC:6179; ITPKB.
DR MIM; 147522; -;
DR GO; GO:0008440; F:inositol-trisphosphate 3-kinase activity; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR005522; IPK.
DR Pfam; PF03770; IPK; 1.
KW Transferase; Kinase; Calmodulin-binding.
FT DOMAIN 768 776 CALMODULIN-BINDING (BY SIMILARITY).
FT POLY-SER.
FT CONFLICT 173 173 R -> H (IN REF. 2).
FT CONFLICT 210 210 P -> S (IN REF. 2).
FT CONFLICT 297 301 GASLT -> ARSEP (IN REF. 2).
FT CONFLICT 408 408 A -> S (IN REF. 2 AND 3).
FT CONFLICT 442 443 RV -> IP (IN REF. 4).
FT CONFLICT 552 552 Q -> P (IN REF. 3).
FT CONFLICT 946 946 AA; 102391 MW; 36C0C74679B1EAID CRC64;
SQ SEQUENCE 946 AA; 102391 MW; 36C0C74679B1EAID CRC64;

Query Match 7.2%; Score 109.5; DB 1; Length 946;
Best Local Similarity 20.3%; Pred. No. 0.22;
Matches 55; Conservative 40; Mismatches 85; Indels 91; Gaps 13;

QY 65 FPPRHGTRLLDTEAQGPSEPHLVLDLLAGFAPCVADIKIGAITW----- 112
DB 711 FVPAYHG-----DVVKDGERYHQ--MDLLADFDSEFCVMDCKMGRTVLEBELTKARKP 763
QY 113 -----PPSSP--EPIYAKLAKRG-----TTSVLLGFRVSGVRVVGPEG 150
DB 764 SLRKDMYOKMIEVDPEATEEKAQAVTKPRYMQWRETISTATLGFRIEIK--KEDG 821
QY 151 AVWRTERPEVKAMDAGVRRVLRVSVSSVADEGMDCALAAVYGGKGVL-----SQREL 206
DB 822 TV---NRDFKTKTREQTEAFRETK-----GNHLLIAYRDLKAI 861
QY 207 KAMFEQTLFHY-----SASILGYDAAVAAAGDGGGVTVKLVDFAH---VAEG----- 254
DB 862 RTTLEVSPPFKCHEVIGSLLFIHDKKEQA-----KVMIDFGKTTLPPEGQTLQHD 913
QY 255 ---DGVIDHNLFLGCLSLKFSIVPEP 281
DB 914 VPMQEGNREDGYLSGLNLLDILTEMSSQDAP 944

RESULT 6
ADD MYCLE
ID ADD MYCLE STANDARD; PRT; 362 AA.
AC O9CCL9; Q49907;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Adenosine deaminase (EC 3.5.4.4) (Adenosine aminohydrolase).
GN ADD OR ML0700 OR L308 C2_206.
OS Mycobacterium leprae
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RA Smith D.R., Robison K.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornaby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT Massive gene decay in the leprosy bacillus."
RL Nature 409:1007-1011(2001).
CC -|- CATALYTIC ACTIVITY: Adenosine + H(2)O = inosine + NH(3).
CC -|- SIMILARITY: Belongs to the adenosine and AMP deaminases family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U00022; AAA17330.1; ALT_INIT.
DR EMBL; AL583919; CAC30209.1; -;
DR PIR; E86996; E86996.
DR HSSP; P03958; 1A4M.
DR Leproma; MLC700; -;
DR HAVAP; MF_00540; -; 1.
DR InterPro; IPR006650; A/AMP deam AS.
DR InterPro; IPR001365; A/AMP deaminase.
DR InterPro; IPR006330; A deaminase.
DR Pfam; PF00962; A deaminase; 1.
DR TIGRFAMs; TIGR01430; aden deam; 1.
DR PROSITE; PS00485; A_DEAMINASE; FALSE NEG.
KW Hydrolase; Nucleotide metabolism; Complete proteome.
FT ACT_SITE 208 208 POTENTIAL.
FT ACT_SITE 267 267 POTENTIAL.
FT ACT_SITE 300 300 POTENTIAL.
FT ACT_SITE 301 301 POTENTIAL.
SQ SEQUENCE 362 AA; 39637 MW; ED0C43B5C92908F5 CRC64;

Query Match 6.4%; Score 97.5; DB 1; Length 362;
Best Local Similarity 26.8%; Pred. No. 0.73;
Matches 56; Conservative 23; Mismatches 91; Indels 39; Gaps 10;

QY 83 EPHEHLVLDLLAGFAP-----CVADIKIGAITWPP--SSPEFYAKLAKDRGT 131
DB 69 EPPSHTV-----AVMQTFEALHRVAYECVEDLAADSVVYAEVRFAPELHIDEGLSFDEV 123
QY 132 TSVLLGFRVSGVRVVGPEG-AVWRTERPEVKAMDAGVRRVLRVSVSSVADEGMDCALAA 190

```

Db 124 ASVLAGEF-ADGERACAEAGNAI--TVRCIVTAMRHAMSRERIAELAIRFRDKGV---VGF 177
QY 191 AVYGGKGG-----VLSQREILKAMFEQTLFFHYSASILLGYDAAAVAGGCGGVT 242
Db 178 DIAGAEAGHPTRHDAFAEYMRNARF---TIHAGEAFGLPSIHEAIAFCGADRILGHG 233
QY 243 VKLVDFAHVAGSDGVDHNFGLGSLIK 271
Db 234 VRIVDDIDVDPGGI-----RLGPLASTIR 258

```

RESULT 7

```

CARA_SULTO
ID CARA_SULTO STANDARD; PRT; 372 AA.
AC Q97008;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Carbamoyl-phosphate synthase small chain (EC 6.3.5.5) (Carbamoyl-
DE phosphate synthetase glutamine chain).
GN CARA OR S11503.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX MEDLINE=21456156; PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Akai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermoacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7."
RL DNA Res. 8:123-140(2001).
CC -!- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +
CC phosphate + L-glutamate + carbamoyl phosphate.
CC -!- PATHWAY: Arginine biosynthesis.
CC -!- PATHWAY: Pyrimidine biosynthesis; first step.
CC -!- SUBUNIT: Composed of two chains; the small (or glutamine) chain
CC promotes the hydrolysis of glutamine to ammonia, which is used by
CC the large (or ammonia) chain to synthesize carbamoyl phosphate (By
CC similarity).
CC -!- SIMILARITY: Belongs to the carA family.
CC -!- SIMILARITY: Contains 1 type-1 glutamine amidotransferase domain.

```

```

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

```

```

EMBL; AF000986; BAB66575.1; -.
DR HAMAP; MF_01209; -.
DR InterPro; IPR006274; CarA synth small.
DR InterPro; IPR001317; CP synthGATase.
DR InterPro; IPR002474; CP_synthsmall.
DR InterPro; IPR000991; GATase.1.
DR Pfam; PF00988; CPase_sm chain; 1.
DR Pfam; PF00117; GATase; 1.
DR PRINTS; PR00099; CPASGATASE.
DR PRINTS; PR00096; GATASE.
DR TIGRfam; TIGR01368; CPaseIIsma1; 1.
DR PROSITE; PS00442; GATASE TYPE 1; 1.
KW Arginine biosynthesis; Pyrimidine biosynthesis; Ligase;
KW Glutamine amidotransferase; Complete proteome.
FT DOMAIN 1 186

```

```

FT DOMAIN 187 372 GLUTAMINE AMIDOTRANSFERASE.
FT ACT_SITE 262 262 GATASE (BY SIMILARITY).
SQ SEQUENCE 372 AA; 41673 MW; 92904F3DCBA3472A CRC64;

```

Query Match 6.1%; Score 93.5; DB 1; Length 372;
 Best Local Similarity 21.7%; Pred. No. 1.6;
 Matches 69; Conservative 43; Mismatches 111; Indels 95; Gaps 15;

```

QY 22 GPLIDSGLFYKPLQAGDGRGEHEVAFYEAFSAHNAVARDTFFPRFHGTRLL---PTE 78
Db 18 GTLIEGCGFGAKGIRAG---EVFTSMNGY---PESLTD---PSYRGQILVITHPLV 66
QY 79 AQPGEPPHVLVDLLAGFQAPCVADIKIGAIT-----WPPSPPEY 120
Db 67 GNYGVPKQKV-EGILTTFES---EQIQVEGLVSVSETDFPKWNSKSLHLLWLGSEVPG 122
QY 121 IA-----KCLAKORGTSVLGLFRVSGVRVVGPGAVNR-----TERPEVKAMD 164
Db 123 LSDVDTRSIVKVRSGVMGVASGVIEDPKKLEKDYDEIDFTQFTSPKAPIVHLGN 182
QY 165 T-----AGVR-----RVLRKY-VSSVADEGMDCALAAAVYGGKGVL 200
Db 183 TGETIVVDGVRHGILYQLHQRGFTIVRVPCKFNVSKI---MDYYPKGWFGNGFGNP 238
QY 201 SQLRELKAMFEQTLFFHYSASILLGYDAAAVAGGCGGVTVKLVDFAHVAGSDGVID- 259
Db 239 NLLQVEIENFKELTEYKIPILGICLGHQVATLWGG-----KYNKKKFGHRAINKPVIDI 293
QY 260 -----HNFLLGLCS 268
Db 294 SSNKCYITTHNHGYGILS 311

```

RESULT 8

```

ID NAGZ_XANCP STANDARD; PRT; 331 AA.
AC Q8PB42;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Beta-hexosaminidase (EC 3.2.1.52) (N-acetyl-beta-glucosaminidase)
DE (Beta-N-acetylhexosaminidase).
GN NAGZ OR XCC1283.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.P.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Fortmighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locall E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Teal S.M., White F.F.,
RA Secubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
CC -!- FUNCTION: Cleaves GlcNAc linked beta-1,4 to MurNAc tripeptides (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of terminal non-reducing N-
CC acetyl-D-hexosamine residues in N-acetyl-beta-D-hexosaminides.

```


Search completed: March 24, 2004, 20:58:10
Job time : 21 secs


```
QY 126 AKDRGTTSVLLGRVSGVRVVGPE-----GAVWRTERPEVKAMDTAGVRRVRLRYVSS-VA 180
Db 122 KKDVESSPFLGRISGLQVYKNDKESSEVVLKPERKLLQMLTADEVRLVKRFVSSNPK 181
QY 181 DEGMDCALAAVYGGKGVLSQRLKAWEEOTLHFYASILLGY--DAAVAAGGDG 238
Db 182 SDQPDSCFAAVYGGSGNIIQAQLLELAKWFEQDTIYHFNCSVLMLKEKTKWNGGEE 241
QY 239 --GGVTYKLVDFAHVAGDGVVDHNFGLGCSLTKFVSDIVPETHPTQP 285
Db 242 SLGCAAVKLIDFAHVTEGNGVDHNFGLGCSLTKFISEILTGPDENSP 290

RESULT 2
ID Q9FLT2 PRELIMINARY; PRT; 300 AA.
AC Q9FLT2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE Similarity to unknown protein (inositol-(1,4,5) triphosphate
DE 3-kinase) (AT5g61760/mac9_60) (inositol polyphosphate 6-/3-/5-kinase
DE 2b).
GN IP3K OR IPK2B.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=98290546; PubMed=9628582;
RA Sato S., Kaneko T., Kotani H., Nakamura Y., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. IV.
RT Sequence features of the regions of 1,456,315 bp covered by nineteen
RT physically assigned P1 and TAC clones."
RL DNA Res. 5:41-54(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Xia H.J., Breatley C.A., Mueller-Roeber B.;
RT "Identification of inositol-(1,4,5) triphosphate 3-kinase from
RT Arabidopsis thaliana."
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Xia H.J., Breatley C., Mueller-Roeber B.;
RT "Identification of inositol-(1,4,5) triphosphate 3-kinase from
RT Arabidopsis thaliana."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Meyers M.C., Banh J., Bowser L.,
RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis ORF clones."
```

```
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RX MEDLINE=22302036; PubMed=12226109;
RA Stevenson-Paulik J., Odom A.R., York J.D.;
RT "Molecular and Biochemical Characterization of Two Plant Inositol
RT Polyphosphate 6-/3-/5-Kinases."
RL J. Biol. Chem. 277:42711-42718(2002).
DR EMBL; AB010069; BAB10076.1; -
DR EMBL; AJ245521; CAC43071.1; -
DR EMBL; AF412073; AAL06526.1; -
DR EMBL; AJ243592; CAC43070.1; -
DR EMBL; AY072621; AAL62012.1; -
DR EMBL; AY147936; AAN63058.1; -
DR GO; GO:0008440; F:inositol-trisphosphate 3-kinase activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR InterPro; IPR005522; IPK.
DR Pfam; PF03770; IPK; 1.
KW Kinase.
SQ SEQUENCE 300 AA; 33487 MW; 6903A3818CEF27D7 CRC64;
Query Match 47.0%; Score 714.5; DB 10; Length 300;
Best Local Similarity 52.8%; Pred. No. 4e-51;
Matches 150; Conservative 39; Mismatches 82; Indels 13; Gaps 8;
QY 4 LHPPEHQVAGHRASAKLQPLDGGSLFYKPLQAGDGRGHEVAFYAFSAHAAVPAIRD 63
Db 2 LKVPHEQVAGHTASDGLGLVDDQGRFFKPLQGDGRGHEAFYFSTSNMKVPOHIH- 60
QY 64 TEPFRPHGTVLLPTEAQFGEPH-PHIVLDDLLAGFOAPCVADIKIGAITWPPSSPPYIA 122
Db 61 RYFPVYHGTVLV--EASDGSGLKPLHLVDDVWVGSGYANPSVMDVKIGSRWYPDVSEYFK 118
QY 123 KCLAKDRGTTSVLLGRVSGVRVVG--PEGAWRTERPEVKAMDTAGVRRVRLRYVSS--SV 179
Db 119 KCIKDKRQTTVSLGRVSGFKTFDHOESSFWAEKKLVGNADGARLARLRFVSSNSP 178
QY 180 ADEGM--DCALAAVYGGKGVLSQRLKAWFEOTLHFYASILLGVDAAAV-AAGG 236
Db 179 ADSNLTPNCAFASVYGGCGNIIQAQLLELKWPEQTQTLHFNSCSILMIYENESILMQGG 238
QY 237 DGG--GVTYKLVDFAHVAGDGVVDHNFGLGCSLTKFVSDIV 277
Db 239 DQAPAPRAQKLVDFAHVLDGNGVDHNFGLGCSLTKFVKDIL 282

RESULT 3
Q8LDQ5 PRELIMINARY; PRT; 300 AA.
AC Q8LDQ5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative inositol hexaphosphate kinase.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation."
RL Genome Biol. 0:0-0(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY085862; AAM63075.1; -
```

```
DR GO: 0008440; F:inositol-trisphosphate 3-kinase activity; IEA.
DR GO: 0016301; F:kinase activity; IEA.
DR InterPro: IPR005522; IPK.
DR Pfam: PF03770; IPK; 1.
KW Kinase.
SQ SEQUENCE 300 AA; 33492 MW; DCB7FC6594BEF1713 CRC64;
Query Match 46.3%; Score 704.5; DB 10; Length 300;
Best Local Similarity 52.1%; Pred. No. 2.7e-50;
Matches 148; Conservative 40; Mismatches 83; Indels 13; Gaps 8;
QY 4 LHPPEHQVAGHRASAKGLPLDGSGLFYKPLQAGDRGEHEVAFYFAFSAHAAVAPARIRD 63
Db 2 LKYPEHQVAGHIASDGLPLVDQGRFFKPLQDGRGEHEAFYFSTNKKVPDHIH- 60
QY 64 TFFPRFHGTRLLPTEAQGPCH-PLHVLDDLAGFOAPCVADIKIGAITWPPSSPEPIAK 123
Db 62 RYFPVYHGTQAV--EGSDGA--AMVLENLALHAYKPSVMDKVGSRTPWPDSEYIQK 117
QY 124 CLAKDRGTTSVLLGFRVSGVRVVG--PEGAVWTERPEVKAMDTAGVRRVRLRYVS--SVA 180
Db 118 CLKKDTGTTTSSGFRISGFVYDHHKESFVKPERKILRGLOVDGARLTILRKVFSSNLS 177
QY 181 DSGM--DCALAAAVYGGKGVLSQLRELKAWPEQTILFHFYASILLGYDAAVAAAGDG 238
Db 178 DTGSKPDSAFASVYGGSHGILTLLELKTWFENQTLVHFNSCSILMIYENESILKGNDD 237
QY 239 GG-VTVKLVDFAHVAGDGVIDHNFGLGCLSLIKFVSDIV 277
Db 238 DARPQVKLVDFAHVLDGNGVIDHNFGLGCLSFINFIREIL 277
RESULT 5
Q9LY23 PRELIMINARY; PRT; 286 AA.
AC Q9LY23;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN T211_80 OR ATSG07370.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,
RA Rudd S., Lemcke K., Mayer K.F.X.;
RA Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RA Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Southwick A., Nguyen M., Tripp M., Palm C.J., Jones T., Wu T.,
RA Carninci P., Chen H., Cheuk R., Chan M.M., Chang C.H., Dale J.M.,
RA Deng J.W., Hayashizaki Y., Hsuan V.W., Lee J.M., Ishida J., Kamiya A.,
RA Kawai J., Kim C.J., Narusaka M., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Shinn P., Tang C.C., Toroumi M., Wallender E.K., Wong C.,
RA Wu H.C., Yamada K., Yu G., Yuan S., Shinozaki K., Ecker J.,
RA Theologis A., Davis R.W.;
RA Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Tripp M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shim P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL163912; CAB87926.1; -
DR EMBL: AY136378; XAM97044.1; -
DR EMBL: BT000196; XAN15515.1; -
DR EMBL: T49876; T49876.
DR GO: 0008440; F:inositol-trisphosphate 3-kinase activity; IEA.
DR InterPro: IPR005522; IPK.
DR Pfam: PF03770; IPK; 1.
KW Hypothetical protein.
SQ SEQUENCE 286 AA; 31946 MW; 98578D433C157EFC CRC64;
Query Match 44.3%; Score 673.5; DB 10; Length 286;
Best Local Similarity 50.4%; Pred. No. 9.5e-48;
Matches 141; Conservative 42; Mismatches 86; Indels 11; Gaps 7;
QY 4 LHPPEHQVAGHRASAKGLPLDGSGLFYKPLQAGDRGEHEVAFYFAFSAHAAVAPARIRD 63
Db 3 LKYPEHQVAGHIAKDGKPGPLVDKGRFFKPLQDGRGEHEAFYFSTNTEVPDHIH- 61
```

```
QY 4 LHPPEHQVAGHRASAKGLPLIDGSLFYKPLQAGDRGEHEVAFYFAFSAHAAPARIRD 63
Db 3 LKYPEHQVAGHIAKDGKPGFLVDKGRFFKPLQGDSSRGIEVKFYFSFSSNTEVPEHIH- 61
QY 64 TFFPRFHGTRLLPTEAQPGEHPHPLVLDLLLAGFQAPCVADIKIGAITWPPSPPEPIAK 123
Db 62 RYFPVYHGTQAV--EGSDGA--AMVLENLLAEYTKPSVMDVKGSRWTWYDASEEYIQK 117
QY 124 CLAKDRGTTSVLLGFRVSGVRVVG--PEGAVWRTERPEVKAMDTAGVRVLRVYVS--SVA 180
Db 118 CLAKDTGTTVSSGFRISGPEVYDHKESFVKPERKLLRGLDVGARLTLRKRVSSNSLS 177
QY 181 DEGM--DCALAAVYGGKGVLSQRLKAWFEEQTLPHFYSASILLGYDAAVAAGDG 238
Db 178 DTGSKPDSAFASVYGGSHGILITQLLELKTWENQTLHFNSCSILMVYENESILKGNDD 237
QY 239 GG-VTVKLVDFAHVAEGDGVIDHNFGLGCLSLIKFVSDIV 277
Db 238 DARPOVKLVDFAHVLDGNGVIDHNFGLGCLSFINFIREIL 277

RESULT 6
Q9ARCI PRELIMINARY; PRT; 285 AA.
AC Q9ARCI;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Linum usitatissimum (flax) (Linseed).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Malpighiales; Linaceae; Linum.
OX NCBI_TaxID=4006;
RN [1]
RP SEQUENCE FROM N.A.
RA Dadds P.N., Lawrence G.J., Ellis J.G.;
RT "Identification of the N rust resistance gene of flax and analysis of
RT the role of intragenic sequence exchange in the evolution of the
RT complex N locus."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ310150; CAC35324.1; -.
DR GO; GO:0008440; F:inositol-trisphosphate 3-kinase activity; IEA.
DR InterPro; IPR005522; IPK.
DR Pfam; PF03770; IPK; 1.
KW Hypothetical protein.
SQ SEQUENCE 285 AA; 31542 MW; 8E6A0514100367BF CRC64;

Query Match 43.7%; Score 664.5; DB 10; Length 285;
Best Local Similarity 50.8%; Pred. No. 5.3e-47;
Matches 141; Conservative 45; Mismatches 83; Indels 13; Gaps 7;

QY 4 LHPPEHQVAGHRASAKGLPLIDGSLFYKPLQAGDRGEHEVAFYFAFSAHAAPARIRD 63
Db 2 LKYPEHQVAGHQDINGLLGPLVDSDGFRFYKPLQDGERGATEMAYFSFNKTIPOXVR- 60
QY 64 TFFPRFHGTRLLPTEAQPGE--HPHLVLDLLLAGFQAPCVADIKIGAITWPPSPPEPIA 122
Db 61 RFFQAFQGTQSI--EASDGSGLPHLITLEDITRLHLCVMDIKIGSRWTWYPEASQAYIE 118
QY 123 KCLAKDRGTTSVLLGFRVSGVRVVG-----PEGAVWRTERPEVKAMDTAGVRVLRVYVS 178
Db 119 KCLAKDVESNPLLGFRISGLQVYNGKESSEVVLKPERKLLNLINADAEVRLVKRFVSS 178
QY 179 -VADEGMDCALAAVYGGKGVLSQRLKAWFEEQTLPHFYSASILLGY--DAAVAAG 235
Db 179 NLKSDQPDCCFAAVVSSGNGILLAQLELKAWEFDQTIHFNSCSVLMLEYEKETKWNG 238
QY 236 GDG--GGVTVKLVDFAHVAEGDGVIDHNFGLGCLSLIKFVSD 275
Db 239 GEESLGAACKLVDFAHVTEGNGVIDHNFGLGCLSLIKFTISE 280
```

```
RESULT 7
Q9LF72 PRELIMINARY; PRT; 286 AA.
AC Q9LF72;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative inositol hexaphosphate kinase (Inositol polyphosphate
DE 6-/3-/5-kinase 2a).
GN IP6K OR IPK2A.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Xu J., Mueller-Roeber B., Xu Z.H., Xue H.W.;
RT "no citation.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22302036; PubMed=12226109;
RA Stevenson-Paulik J., Odom A.R., York J.D.;
RT "Molecular and Biochemical Characterization of Two Plant Inositol
RT Polyphosphate 6-/3-/5-Kinases.";
RL J. Biol. Chem. 277:42711-42718(2002).
DR EMBL; AJ404678; CAB96043.1; -.
DR EMBL; AY147935; AAN63057.1; -.
DR GO; GO:0008440; F:inositol-trisphosphate 3-kinase activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR InterPro; IPR005522; IPK.
DR Pfam; PF03770; IPK; 1.
KW Kinase.
SQ SEQUENCE 286 AA; 31960 MW; 83C47D433C0480FE CRC64;

Query Match 43.3%; Score 658.5; DB 10; Length 286;
Best Local Similarity 49.6%; Pred. No. 1.7e-46;
Matches 139; Conservative 42; Mismatches 88; Indels 11; Gaps 7;

QY 4 LHPPEHQVAGHRASAKGLPLIDGSLFYKPLQAGDRGEHEVAFYFAFSAHAAPARIRD 63
Db 3 LKYPEHQVAGHIAKDGKPGFLVDKGRFFKPLQGDSSRGIEVKFYFSFSSNTEVPEHIH- 61
QY 64 TFFPRFHGTRLLPTEAQPGEHPHPLVLDLLLAGFQAPCVADIKIGAITWPPSPPEPIAK 123
Db 62 RYFPVYHGTQAV--EGSDGA--AMVLENLLAEYTKPSVMDVKGSRWTWYDASEEYIQK 117
QY 124 CLAKDRGTTSVLLGFRVSGVRVVG--PEGAVWRTERPEVKAMDTAGVRVLRVYVS--SVA 180
Db 118 CLAKDTGTTVSSGFRISGFRISGFRISGFRISGFRISGFRISGFRISGFRISGFRISGFRIS 177
QY 181 DEGM--DCALAAVYGGKGVLSQRLKAWFEEQTLPHFYSASILLGYDAAVAAGDG 238
Db 178 DTGSKPDSAFASVYGGSHGILITQLLELKTWENQTLHFNSCSILMVYENESILKGNDD 237
QY 239 GG-VTVKLVDFAHVAEGDGVIDHNFGLGCLSLIKFVSDIV 277
Db 238 DARPOVKLVDFAHVLDGNGVIDHNFGLGCLSFINFIREIL 277

RESULT 8
Q9VPR6 PRELIMINARY; PRT; 309 AA.
AC Q9VPR6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG13688 protein.
GN CG13688.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
```


[4]
RN SEQUENCE FROM N.A.
RP Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
[5]
RN SEQUENCE FROM N.A.
RP FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AEO03589; AAF51477.2; --
DR FlyBase: FBgn0031267; CG13688.
DR GO: GO:0008440; Farnesitol-trisphosphate 3-kinase activity; IEA.
DR InterPro: IPR005522; IPK.
DR Pfam: PF03770; IPK; 1.
SQ SEQUENCE 309 AA; 35330 MW; 3202DDF886094912 CRC64;

Query Match 18.9%; Score 288; DB 5; Length 309;
Best Local Similarity 30.1%; Pred. No. 9.4e-16;
Matches 97; Conservative 47; Mismatches 114; Indels 64; Gaps 14

QY	2	SDIHPPPE-----HOVAGH---RASASKGLPIIDG-SGLFYFKPLOGDGRGEHEVAIFYEAF	51
Db	4	SQELPEGFROLTKTVAGHTFEESNAEAVGLIQDSKAGCVLKPLGKPCGERELRFYESL	63
QY	52	SAHA-----VPARTDTFFPHGTTRLPTPAQGPHEPHLVLDLLIAGFAQCVDADIK	106
Db	64	AEAGASGDNDLLALLRG-HVPRFYGPLKLIVNRR---ERTFLRLDLTRSYAKPCVMGVK	119
QY	107	IGAITW-PPSSP-----EPYIAKCLAKDRGTSTVLGLFRYSVGVRVVGSP	153
Db	120	MGRRTWDPESSNKRVBEAKYVMCKQK-----LGLCLPGFGVYLPKSEHTQTETIL	171
QY	154	RTERPEVKAMDTAGVRRVLRVSVSSVADEGM-----CALAAAVYGKGGLVSQRELKAW	209
Db	172	RHKDYKGSINVVEGFKTMALEFNASTSDSKSRACELLK-----EVLRLQLQEILLAW	225
QY	210	FREOTLFHYSSAILLYDAAVA-----AGSGGGVTVKLVDPFAHVAEGD-	255
Db	226	FQRQLLHFYASSLLICYDSLAPDKPKPLINGYHONDDPATWRVWKMDIFARVYPAEQ	285
QY	256	GVIDHNFLGGLCSLIKFSVDIV 277	
Db	286	GLPDENYMFGLQSLEVVQSIL 307	

RESULT 9
QBMRX9 PRELIMINARY; PRT; 309 AA.

ID	QBMRX9	PRELIMINARY; PRT; 309 AA.
AC	QBMRX9;	
DT	01-OCT-2002 (TREMBLrel. 22, Created)	
DT	01-OCT-2002 (TREMBLrel. 22, Last sequence update)	
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)	
DE	SD14726P.	
GN	CG13688.	
OS	Drosophila melanogaster (Fruit fly).	
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;	
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	
OC	Nebriidae; Drosophilidae; Drosophila.	
OX	NCBI_TaxID=7227;	
RN	SEQUENCE FROM N.A.	
RP	Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,	
RA	Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,	
RA	George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,	
RA	Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,	
RA	Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,	
RA	Celniker S.;	
RL	Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.	
EMBL	AY119207; AAM51067.1; --	
DR	FlyBase: FBgn0031267; CG13688.	
DR	GO: GO:0008440; Farnesitol-trisphosphate 3-kinase activity; IEA.	
DR	InterPro: IPR005522; IPK.	
DR	Pfam: PF03770; IPK; 1.	
SQ	SEQUENCE 309 AA; 35312 MW; 7EA209AC834C5EAF CRC64;	

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Rhydroidea; Drosophilidae; Drosophila.
[1]
RN NCBI_TaxID=7227;
RP SEQUENCE FROM N.A.
RC MEDLINE=Berkley;
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anantides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Held G., Nelson C.R., Miklos G.L.G.,
RA Abriel J.F., Agbayani A., An H.-J., Andrews-Fiankoch C., Baldwin D.,
RA Balow R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Beriman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foeller C., Gabrielian A.E., Gang N.S., Galbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Kaulous K.A., Howland T.J., Wei M.-H., Ibegwan K.,
RA Jalali M., Koulton F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Jalali M., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
[2]
RN SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Anantides P.G., Brandon R.C., Rogers Y.,
RA Hanson J., An H., Baldwin D., Banon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwan K., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskaas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
[3]
RN SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradscky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Bernan B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.


```
Q8NFUS PRELIMINARY; PRT; 416 AA.
AC Q8NFUS;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Inositol polyphosphate multikinase.
GN IPMK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22176314; PubMed=12027805;
RA Nalaekowski M.M., Deschermeier C., Panick W., Mayr G.W.;
RT "The human homologue of yeast ArgR111 protein is an inositol phosphate
RT multikinase with predominantly nuclear localization.";
RL Biochem. J. 366:549-556(2002).
DR EMBL; AF432853; AAM97838.1; -
DR GO; GO:0008440; F:inositol-trisphosphate 3-kinase activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR InterPro; IPR005522; IPK.
DR Pfam; PF03770; IPK; 1.
KW Kinase.
SQ SEQUENCE 416 AA; 47222 MW; 622D678696A892F9 CRC64;

Query Match 13.2%; Score 201.5; DB 4; Length 416;
Best Local Similarity 28.4%; Pred. No. 2e-08;
Matches 71; Conservative 34; Mismatches 110; Indels 35; Gaps 9;

QY 6 PPEHQVAGHRASASKLGLPLDGLSGFYKPLQAGDRGEHEVAFYE---AFSAHAAVPAIR 62
D 48 PLSHQVAGHMYGKDKVGLIQLHPDGVTLKQLQPPRGPRLFEFYMVYAADCFDGVLELR 107
QY 63 DTFPRFHGTRLLPTEAQGEHPHVLVDLLAGFOAPCVADIKIGAITWPPSSPEFYIA 122
D 108 -KYLPRYGIWSPPT--APND--LVLKEDVTHKFNKPCIMDVKIGQKSDYPPASSEKIQ 162
QY 123 KCLADRGTTSVLLGPRVSGVRVVGEGAVWTE-----RPEVKAMDTAGVRVLRYS 177
D 163 QQVSKYPLMEEI--GFLVLGMRYVHVSYSYETENOHYGRSITKETIKDGVSRFFHNGYC 220
QY 178 SVADGMDCALAAVYGGKGVLSQLRELKAWFEEQTLFHEYSASILLGYDAAAAGD 237
D 221 LRKD-----AAVASI-----QKTEKTLQWFENKQLNFYASSLLFYE-----GS 260
QY 238 GSGVTIVKLVD 247
D 261 SQFTTKLND 270

RESULT 13
Q8BZA8 PRELIMINARY; PRT; 371 AA.
AC Q8BZA8;
DT 01-WAR-2003 (TrEMBLrel. 23, Created)
DT 01-WAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Inositol polyphosphate MULTIKINASE homolog (Fragment).
GN IPMK OR 2410017C19IRK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
```

```
RL Nature 420:563-573(2002).
DR EMBL; AK036084; BAC29300.1; -.
DR EIR; PT0645; PT0645.
DR MGD; MGI:1916968; Impk.
DR GO; GO:0008440; F:inositol-trisphosphate 3-kinase activity; IEA.
DR InterPro; IPR005522; IPK.
DR Pfam; PF03770; IPK; 1.
FT NON_TER 1
SQ SEQUENCE 371 AA; 41888 MW; 1C9529B05DA34DF5 CRC64;

Query Match 13.1%; Score 199.5; DB 11; Length 371;
Best Local Similarity 26.3%; Pred. No. 2.6e-08;
Matches 70; Conservative 47; Mismatches 122; Indels 27; Gaps 10;

QY 6 PPEHQVAGHRASASKLGLPLDGLSGFYKPLQAGDRGEHEVAFYEAFSAHAAVPAIRD-- 63
D 6 PLSHQVAGHMYGKDKVGLIQLHPDGVTLKQLQPPRGPRLFEFYMVYAADCFDGVLELR 65
QY 64 TFFPRFHGTRLLPTEAQGEHPHVLVDLLAGFOAPCVADIKIGAITWPPSSPEFYIAK 123
D 66 KHLPRYGIWSPPT--APND--LVLKEDVTHKFNKPCIMDVKIGKSDYPPASSEKIQQ 121
QY 124 CLAKDRGTTSVLLGPRVSGVRVVGEGAVWTEPEVKAMDTAGVRVLRVYVSSVADEG 183
D 122 QVSKYPLMEEI--GFLVLGMRYVHLSDSYETQNHGVRGLT---KETLKEGYSKFFHNG 176
QY 184 MDC---ALAAVYGGKGVLSQLRELKAWFEEQTLFHEYSASILLGYDAAAAGDGG 239
D 177 F-CLRKDAVASI-----QKTEKTLQWFENKQLNFYASSLLFYEGSQPATTKAN 227
QY 240 GVTV--KLVDFAHVAEGDGV-IDHNF 262
D 228 DRTLGRFLSKPLTDADGLECNMF 253

RESULT 14
Q7TT16 PRELIMINARY; PRT; 396 AA.
AC Q7TT16;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 27, 2004, 05:59:59 ; Search time 3160 Seconds
(without alignments)
2731.067 Million cell updates/sec

Title: US-10-042-894A-8
Perfect score: 1521
Sequence: 1 MSDLLPHEQVAGHRASAK.....IKFVSIVPETHPTQLGPS 289

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US10042894/runat_24032004_145047_16959/app_query.fasta_1.455
-DB=EST_QPWT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOCPL=0 -LOCPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptp -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10042894 @CGN 1.1.3437 @runat_24032004_145047_16959 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=10 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rtd.*
26: em_gss_pbg.*
27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	1521	100.0	960	28	CC344833 OGIAQ20TV
2	1406.5	92.5	923	29	CG240397 OGYCI27TV
3	1256	82.6	818	29	CG453578 OG9AD40TV
C	1248	82.1	873	28	CC344824 OGIAQ20TV
4	1183	77.8	867	29	CG724960 OGLAB23TV
5	1149	75.5	719	29	CG284572 OGWIF61TV
6	1123	73.8	801	13	CA134480 SCJFRT106
7	1034	68.0	656	14	CA245555 SCBGFL508
8	993.5	65.3	722	13	CA130685 SCCRT100
9	992	65.2	591	29	CG776236 1123005E0
10	987	64.9	671	14	CA202412 SCRLFL100
11	981	64.5	3374	11	AY109355 Zea mays
12	976	64.2	722	29	CC724950 OGAB23TH
C	932	61.3	744	13	CA130686 SCCRT100
13	926	60.9	1191	14	CK211413 FGAS02325
C	921.5	60.6	593	28	BZ774817 i151a11b
14	920	60.5	960	29	CG073495 PUCG54TD
C	903	59.4	955	28	CC384766 FUHOG53TD
15	869.5	57.2	794	10	BF617713 HVSMB001
C	811.5	53.4	907	29	CG240387 OGYCI27TH
16	807	53.1	575	29	CG734025 1119161G0
17	770	50.6	619	14	CA204213 SCAGFL109
C	758	49.8	1099	14	CK215908 FGAS02788
C	724.5	47.6	1123	14	CK214494 FGAS02642
18	724	47.6	910	14	BH408067 1007056A0
19	714	46.9	551	28	BH408067 1007056A0
20	644	42.3	557	10	BF277833 GA_Eb003
C	642	42.2	831	29	CG953151 MBFEF66TF
21	640	42.1	620	13	BQ620191 Talrl176F
22	613	40.3	453	12	B1956405 HVSMB001
23	606	39.8	433	12	EM099879 EBS01_SQ
24	584	38.4	448	28	CC028437 3591_1_10
25	581.5	38.2	478	12	BG417083 HVSMB001
26	580	38.1	876	29	CG073494 PUCG54TB
27	567	37.3	409	12	EM099878 EBS01_SQ
C	562	36.9	927	29	CG284580 OGWIF61TV
28	559	36.8	431	29	CG776505 1123005E0
C	529	34.8	317	29	CG194403 PUIGW38TB
29	527	34.6	767	29	CG971224 MBFEF51TF
C	527	34.6	823	29	CG931827 MBFCW31TF
30	527	34.3	688	13	BQ405077 GA_Ed007
31	517	34.0	739	14	CA917416 EST641563
32	508.5	33.4	629	13	BUR93995 HM0521R
C	496	32.6	691	28	BH585814 BOGHV81TF

ALIGNMENTS

RESULT 1
CC344833
LOCUS OGIAQ20TV ZM 0.7.1.5 KB Zea mays genomic clone ZMMBMA0359C15, 960 bp DNA linear GSS 16-MAY-2003
DEFINITION genomic survey sequence.
ACCESSION CC344833
VERSION CC344833.1 GI:30814239
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.


```
QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20
Db 34 ATGCCCGACCTCCACCCGCGGAGCACCAAGTCGCGGTCACCGGCTCCGCCAGCAG 93
QY 21 LeuGlyProLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
Db 94 CTGGGCGCGCTCATCGACGCGCTCGCGCTCTCTCAAGCGGCTCCAGGCGCGACCGT 153
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArg 60
Db 154 GGGGAGCAGAGTCGCTCTATAGAGCGGTCTCCGCCACGCGCGCTCCGCCGCGC 213
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 214 ATCCGAGACACCTCTCTCCCGCGTTCACGCGACCGGACTCTCCGCCACCGGCGAG 273
QY 81 ProGlyGluProHisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaPro 100
Db 274 CCGGGGAGCGCATCCGACCTCGTCTCGACGACCTCTCGCGGCTTCAGGCGGCC 333
QY 101 CysValAlaAspTyrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120
Db 334 TCGCTCGCAGACATCAAGATCGCGCATCAGTGGCCACCGAGTTCGCGGAGCGCTAC 393
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 394 ATCCCAAGTACTCGCCCAAGGACCGCGGACCGGAGCGGCTCTCGGATTCGCGT 453
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTyrArgThrGluArgProGluVal 160
Db 454 TTGC--GTCCGAGTCGTCGCGCGCGCGGCGCGTGTGGCGGACGCGCGCGGAGTG 511
QY 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSer-ValAl 180
Db 512 AAGGCTATGGACACCGTTCGCGCGTTCGCGCGTTCGCGCGTTCGCGCGTTCGCGT 571
QY 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyValle 200
Db 572 CGACGAGGGATGGACTGCGCGCTCGCGCGCGGTGTACGAGGAGAAAGTGGAGTCT 631
QY 200 uSerGlnLeuArgGluLeuLysAlaTyrPheGluGluGlnThrLeuPheHisPheTyrSe 220
Db 632 GTACAGCTGCGGAGCTCAAGGCTATGGTGGAGGAGCAGCTCTGTTCCACTTCTACT 691
QY 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyGlyGly 240
Db 692 GCGCTCGATCTCTGGGCTATGATGCTGCTCAGTCGAGCGGAGCGGAGGTGGGCGTG 751
QY 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260
Db 752 GGTACAGTGAAGCTGGTGGACTTTGCCCATGTGGCGAGGCTGATGGGTTGATTGACCA 811
QY 260 sAnPheLeuGlyGlyLeuCysSerLeuIleLysPheValSerAspIleValProGluTh 280
Db 812 CAACTTCCTGGCGAGCTCTGCTAGTGTATCAAGTTCTGTCATTTGTCATTTTCCAGAG 871
QY 280 rProHisThrGlnProLeuGlyProSer 289
Db 872 TCTTAGCGAGCTTTGGGTCCTTCT 899

RESULT 3
CG453578
LOCUS
DEFINITION
  OG9AD40TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0849H08,
  genomic survey sequence.
ACCESSION
  CG453578
VERSION
  CG453578.1 GI:34838578
KEYWORDS
  GSS.
SOURCE
  Zea mays
  ORGANISM
    Zea mays
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
    clade; Panicoideae; Andropogoneae; Zea.
```

```
REFERENCE 1 (bases 1 to 818)
  Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
  Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
  Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  Consortium for Maize Genomics
  Unpublished (2002)
  Other GSSs: OG9AD40TH
  Contact: Cathy Whitelaw

TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA
  Tel: 301-838-5843
  Fax: 301-838-0208
  Email: whitelaw@igr.org
  Seq primer: TF
  Class: sheared ends.

FEATURES
  source
    1..818
    /organism="Zea mays"
    /mol_type="genomic DNA"
    /strain="B73"
    /db_xref="taxon:4577"
    /clone="ZMMBMA0849H08"
    /clone_lib="ZM 0.7 1.5 KB"
    /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
    methylation filtered genomic DNA library"

ORIGIN
Alignment Scores:
  Pred. No.:      5,92e-112      Length:      818
  Score:          1256.00        Matches:    239
  Percent Similarity: 100.00%    Conservative: 0
  Best Local Similarity: 100.00%  Mismatches: 0
  Query Match:      82.58%       Indels:     0
  DB:               29          Gaps:         0
```

US-10-042-894a-8 (1-289) x CG453578 (1-818)

```
QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20
Db 96 ATGTCGACCTCCACCCGCGGAGCACCAAGTCGCGCGCGCTCCGCCAGCAG 155
QY 21 LeuGlyProLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
Db 156 CTGGGCGCGCTCATCGACGCGCTCGCGCTCTCTCAAGCGCTCCAGGCGCGACCGT 215
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaAlaValProAlaArg 60
Db 216 GGGGAGCAGAGTTCGCTCTATGAGCGGTTCTCGGCGGCTTCGCGCGCGCGCGC 275
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 276 ATCCGAGACACCTCTCTCCCGGTTCCAGGCGGAGCTCTCCGCCACCGGCGGCGAG 335
QY 81 ProGlyGluProHisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaPro 100
Db 336 CCGGGGAGCGCATCTCTACCTCGTCTCGACGACCTCTCCGCCGCGGTTTCAGGCGGCC 395
QY 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProSerSerProGluProTyr 120
Db 396 TGCCTCGCAGACATCAAGATCGGCGGCTTCACGTCGCGGCGGCGGAGTTCGCGGAGCC 455
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 456 ATCCGAGAGTCCCTCGCCAGAGACCGCGGAGCGGCGGCTGTGGCGGAGCGGAGGTG 515
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTyrArgThrGluArgProGluVal 160
Db 516 TCCGCGCTCCGAGTCGTTCGCGCGCGCGGCGGCGGCTGTGGCGGAGCGGAGGTG 575
QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgValLeuArgTyrValSerSerValAla 180
Db 576 AAGGCTATGGACACCGCGCGGCTTCGCGCGGCTTCGCGCGGCTTCGCTCATCGTTGCC 635
```

```

QY 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyValLeu 200
Db 636 GAGCAGGGATGACTGTGGCTCGCGCGCGGTGTACGAGGAAAAGGTGGAGTCTTG 695
QY 201 SerGlnLeuArgGluLeuLysAlaTTPheGluGluGlnThrLeuPheHisPheTyrSer 220
Db 696 TCACAGCTGGCGAGCTCAAGGGTGGTTCAGAGAGCAGACTCTGTCCACTTCTACTCG 755
QY 221 AlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyAspGlyGly 239
Db 756 GCCTCGATTCTTCTGGCTATGATGCTGTGCAAGTCCGACAGCGGAGATGGGGG 812

RESULT 4
CC344824/c 873 bp DNA linear GSS 16-MAY-2003
LOCUS OGIAQ20TH ZM 0.7.1.5.KB Zea mays genomic clone ZMBMa0359C15,
DEFINITION genomic survey sequence.
ACCESSION CC344824
VERSION CC344824.1
KEYWORDS GI:30814230
SOURCE GSS.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1. (bases 1 to 873)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
source
1. .873
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMBMa0359C15"
/clone_lib="ZM 0.7.1.5.KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Alignment Scores:
Pred. No.: 3,94e-111 Length: 873
Score: 1248.00 Matches: 238
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 82.05% Indels: 0
DB: 28 Gaps: 0

US-10-042-894A-8 (1-289) x CC344824 (1-873)
QY 52 SerAlaHisAlaValProAlaAlaArgIleArgAspThrPhePheProArgPheHisGly 71
Db 872 TCCGCCACACCGCGCTCCCGCGCGCATCCGACACCTTCTTCCCGCGGTCCACGCG 813
QY 72 ThrArgLeuLeuProThrGluAlaGlnProGlyGluProHisProHisLeuValLeuAsp 91
Db 812 ACCGACTCTCTCCACCGAGGCGCAGCCCGGAGCGCGATCTCTACCTCGTCTCGAC 753
QY 92 AspLeuLeuAlaGlyPheGlnAlaProCysValAlaAspIleLysIleGlyAlaIleThr 111
Db 752 GACTCTCTCGGGGTTTCAGCGCGCTCGTGCAGACATCAAGATCGCGCCATCACG 693

```

```

QY 112 TrpProProSerSerProGluProTyrIleAlaLysCysLeuAlaLysAspArgGlyThr 131
Db 692 TGGCCACCCAGATTCCGCGAGGCCCTACATCGCAAGTGCCTGCCAAGGACCCGCGGACC 633
QY 132 ThrSerValLeuLeuGlyPheArgValSerGlyValArgValValGlyProGluGlyValAla 151
Db 632 ACAGCGTTCTCTCGTGGATTCCGCGTCTCCGGGTCCTCCAGTGTGTGGCCCCCGAGGGCGCC 573
QY 152 ValTrpArgThrGluArgProGluValLysAlaMetAspThrAlaGlyValArgArgVal 171
Db 572 GTGTGGCGACGAGCGCCCGGAGGTGAAGGCCATGGACACCGCGCGCTCCGCGCGCTG 513
QY 172 LeuArgArgTyrValSerSerValAlaAspGluGlyMetAspCysAlaLeuAlaAla 191
Db 512 CTCGCGCGCTAGCTGTTCATCCGTTCGCGACGAGGGATGACTGTGTGCTCGCGCGCG 453
QY 192 ValTyrGlyGlyLysGlyGlyValLeuSerGlnLeuArgGluLeuLysAlaTrpPheGlu 211
Db 452 GTGTACGAGGAAAAGGTGGAGTCTTGTTCACAGCTCCGCGAGCTCAAGCGGTGGTTCGAG 393
QY 212 GluGlnThrLeuPheHisPheTyrSerAlaSerIleLeuLeuGlyTyrAspAlaAla 231
Db 392 GAGCAGACTCTCTTCCACTTCTACTCGCGTGCATTTCTTCTGGCTATCATGCTGCTGCA 333
QY 232 ValAlaAlaGlyGlyAspGlyGlyValThrValLysLeuValAspPheAlaHisVal 251
Db 332 GTCGACAGCGGAGGATGGGGTGGGTGAGCGGTGAGCTGAGCTGCGCATGTTGCCCATGTG 273
QY 252 AlaGluGlyAspGlyValIleAspHisAsnPheLeuGlyGlyLeuCysSerLeuIleLys 271
Db 272 GCGAGGGTGTGGGTGATTGACCAACTTCTTGGGGGGGCTCTGCTGCTGATCAAG 213
QY 272 PheValSerAspIleValProGluThrProHisThrGlnProLeuGlyProSer 289
Db 212 TTCTGTTCTGACATTGTCGCGAGACTCTCTATACGACGCTTGGTCTCTTCT 159

RESULT 5
CC724960 867 bp DNA linear GSS 23-JUN-2003
LOCUS OGLAB23TV ZM 0.7.1.5.KB Zea mays genomic clone ZMBMa0307C22,
DEFINITION genomic survey sequence.
ACCESSION CC724960
VERSION CC724960.1
KEYWORDS GI:32143893
SOURCE GSS.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1. (bases 1 to 867)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGLAB23TH
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
FEATURES
source
1. .867
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMBMa0307C22"
/clone_lib="ZM 0.7.1.5.KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb"

```


methylation filtered genomic DNA library"

ORIGIN

Alignment Scores:

Pred. No.: 8.54e-105 Length: 867
 Score: 1183.00 Matches: 236
 Percent Similarity: 90.15% Conservative: 2
 Best Local Similarity: 89.39% Mismatches: 7
 Query Match: 77.78% Indels: 20
 DB: 29 Gaps: 1

US-10-042-894A-8 (1-289) x CC724960 (1-867)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20
 Db 131 ATGCCGACCTCCACCGCGGAGCACCAAGTCGCGGTCAACGCGCTCCGCCGACGAAAG 190
 QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
 Db 191 CTGGGCCACTCATCGACGACTCTGGCTCTTCTACAGCGCTCCAGGCGCGGACCGT 250
 QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
 Db 251 GGGGAGCAGCGCTCTATAGCGCTTCTCGCCACCGCGCGCTCCGCGCGCGC 310
 QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
 Db 311 ATCCGAGACACTTCTTCCCGCGGTTCACGCGACGACTCTCTCCACCGAGGCGAG 370
 QY 81 ProGlyGluProHisProHisLeuValLeuAspLeuAlaGlyPheGlnAlaPro 100
 Db 371 CCGCGGAGCGCATCCGACCTCTCTCGACGACTCTCTCGGGGTTTGAGGCGCC 430
 QY 101 CysValAlaAspIleIleGlyAlaIleThrTrpProSerSerProGluProTyr 120
 Db 431 TCGCTCGCAGACATCAAGATCGGTGCCATCACGTG----- 465
 QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
 Db 466 -----ACCACGAGCGTTCTGCTCGGATTCCGCGTC 495
 QY 141 SerGlyValArgValValGlyProGlyAlaValTyrArgThrGluArgProGluVal 160
 Db 496 TCGCGGTCGAGTCGCGCGCGGAGGCGCGTGTGCGGACGAGCGCGCGGAGTG 555
 QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSer-ValAl 180
 Db 556 AAGGCTATGGACATTCGCGCGTCCGCGGTCTCTCGCGCTTACGTGTATCCGCTGC 615
 QY 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyGlyValle 200
 Db 616 CGACGAGGGGATGGACTGCGCTCGCGCGCGGTGTACGGAGGAAAGGTGGAGTCTT 675
 QY 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220
 Db 676 GTCACAGCTGCGGAGCTCAAGSCGTGTTGAGGGGCGAGCTCTCTTCCACTTCTACTC 735
 QY 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGly 240
 Db 736 GCGCTCGATTCTTCTGGGTATGATGCTGTGAGTCGAGTCGACGAGCGAGGTGGGGTGG 795
 QY 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValleAspHi 260
 Db 796 GGTAAACAGTGAAGCTGGTGGACTTTCCTCATGTGGCGAGGGGTGATGGGTGATTACCA 855
 QY 260 sAenPheLeu 263
 Db 856 CAATCTCTTG 865

RESULT 6

CG284572

LOCUS

DEFINITION OGWIF61TH ZM_0.7_1.5_KB Zea mays genomic clone ZMBM0584K02,

genomic survey sequence.

CG284572
 CG284572.1 GI:34198786
 GSS.
 Zea mays
 Zea mays

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 719)

REFERENCE

AUTHORS

Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,

Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,

Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

Consortium for Maize Genomics

Unpublished (2002)

Other GSSs: OGWIF61TV

Contact: Cathy Whitelaw

TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TR

Class: sheared ends.

FEATURES

source

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMBM0584K02"

/clone_lib="ZM_0.7_1.5_KB"

/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb

methylation filtered genomic DNA library"

ORIGIN

Alignment Scores:

Pred. No.: 1.33e-101 Length: 719
 Score: 1149.00 Matches: 230
 Percent Similarity: 89.92% Conservative: 2
 Best Local Similarity: 89.15% Mismatches: 7
 Query Match: 75.54% Indels: 20
 DB: 29 Gaps: 1

US-10-042-894A-8 (1-289) x CG284572 (1-719)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20
 Db 3 ATGCCGACCTCCACCGCGGAGCACCAAGTCGCGGTCAACGCGCTCCGCCGACGAAAG 62
 QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
 Db 63 CTGGGCGGACCTCATCGACGACTCTGGCTCTTCTACAGCGCTCTCCAGCGCGGACCGT 122
 QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
 Db 123 GGGGAGCAGAGGTGCGCTTCTATGAGCGGTCTCCGCCACCGCGCTCCGCGCGCGC 182
 QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
 Db 183 ATCCGAGACACCTTCTTCCCGGTTCCACGCGACGCGACTCTTCCCGACCGGCGGAG 242
 QY 81 ProGlyGluProHisProHisLeuValLeuAspLeuAspLeuAlaGlyPheGlnAlaPro 100
 Db 243 CCGGGGAGCGCATCCGACCTCTCTCTCGACGACCTCTCTCGGGGTTTGAGGCGGCC 302
 QY 101 CysValAlaAspIleIleGlyAlaIleThrTrpProSerSerProGluProTyr 120
 Db 303 TCGCTCGCAGACATCAAGATCGGTGCCATCACGTG----- 337
 QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
 Db 338 -----ACCACGAGGTTCTGCTCGGATTCGCGTTC 367

```

QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
Db 368 TCCGGCGTCAGTTCGCGCGCCGAGGCGCCGTGGCGACGAGCGCCGAGGTG 427
QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgArgTyrValSerSer-ValAl 180
Db 428 AAGGCTATGACATGTGCGCGTCCGCGCGTGTCCGGCGTACGTGTATCCGCTTGC 487
QY 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyGlyVal 200
Db 488 CGACAGGGGATGACTGCGCGTCCGCGCGCGTGTACGAGGAGGAGGTGAGTCTT 547
QY 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220
Db 548 GTACAGCTGGCGAGCTCAAGGCGTGTGTCGAGGGGCGAGCTGTGTCCACTTCTACTC 607
QY 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyVal 240
Db 608 GGCCTCGATCTTCTGGGCTATGATGCTGTGCAATGTCAGAGCGGAGGTGGGGTGG 667
QY 240 YValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyVal 257
Db 668 GSTAACAGTGAAGCTGTGGACTTGGCCATGTGCGCGGAGGTGATGGGGTG 719

RESULT 7
LOCUS CAL134480 801 bp mRNA linear EST 24-SEP-2003
DEFINITION SCUFRT1061H1.1 g RT1 Saccharum officinarum cDNA clone SCUFRT1061H1
5', mRNA sequence.
ACCESSION CAL134480
VERSION CAL134480.1 GI:35021536
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum.
1 (bases 1 to 801)
Vettore,A.L., da Silva,F.R., Kemper,E.I. and Arruda,P.
The libraries that made SUCST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 061 row: H column: 11
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1 .801
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCUFRT1061H1"
/lab_host="DH10B"
/clone_lib="RT1"
/note="Organ: Root tips (0.3cm-long) from adult plants;
Vector: pSport1; Site1: SalI; Site2: NotI; An
unidirectional cDNA library generated from [Root tips
(0.3cm-long) from adult plants]. cDNA was prepared from
polyA+ mRNA using SuperScript Plasmid System Kit
(Invitrogen). The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 Kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
http://sucst.lad.ic.unicamp.br/public"

```

ORIGIN

Alignment Scores: 5.39e-99 Length: 801
Pred. No.: 1123.00 Matches: 220
Score: 89.58% Conservative: 12
Best Local Similarity: 84.94% Mismatches: 24
Query Match: 73.83% Indels: 3
DB: 13 Gaps: 0

US-10-042-894A-8 (1-289) x CA134480 (1-801)

```

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20
Db 31 ATGTCGACCTCCGCCCGCGGAGCACCAGTCCGCGGCCACCGCGCTCCGCCAACAG 90
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
Db 91 CTGGGTCCGCTCATCGACGGCTCTGGCTCTTCTACAGCGCTCCAGCGCTCGGGACCGC 150
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
Db 151 GGGGAGCAGCAGCTCGCTTCTACGAGCGGTTCACACCCACCGCGCTCCGCGCGCGC 210
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 211 ATCCGGGACACCTCTTCCCGCGGTTCACGCGCACGCACTCTCTCCACCGAGCGCGG 270
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100
Db 271 CCAGGGGAGCGCATCCGCACTCTCGCTCTCGACGACTCTCTCGCGGGCTGGAGCGCC 330
QY 101 CysValAlaAspIleLysIleGlyAlaIleThr-TrpProProSerSer-ProGluProT 120
Db 331 TGGCTCGCGACATCAAGATCGCGCCATCACAGTGGCGCGCCGAGCTCGCGCAGAGCCCT 390
QY 120 YrIleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgV 140
Db 391 ACCTGTCAAGTCCCTGGTCAAGGACCGCGGACACGAGCATTCCTCGGATTCGCGG 450
QY 140 alSerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluV 160
Db 451 TCTCCGGCTCCGGGTCTCGCTCCGAGGCGCGCGCTGTGGCGACTGACGCGCGGAGG 510
QY 160 alLysAlaMetAspThrAlaGlyValArgArgValLeuArgTyrValSerSerValA 180
Db 511 TGAAGCCCTGGACACCGCGCGTCCGCGCTGTCTCCGCGCTACGTTTCATCCGTTG 570
QY 180 laAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyVal 200
Db 571 CCAGCAGGGGATGGACTGCGCACTCCCGCGCGGTGTACGCGGGAAGAGGAGTTT 630
QY 200 euSerGlnLeuArgGluLeuLysAlaTrpPheGluGlnThrLeuPheHisPheTyrS 220
Db 631 TGTCAACAATTGGCGAGCTTAAGCGCTGGTTTTCGAGGACACAGACTCTGTGTCACCTTACT 690
QY 220 erAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGly-AspGlyGly 239
Db 691 CGGGGGCGATTTCTTCTGGGCTATGAAGTAGTCCGCTTGCACACCGGAGGTAAGGTT 750
QY 240 GlyValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGly 256
Db 751 GGGGTGAAGTAAACTGGGGGCCCTTTCGCAATGTGGCCCAAGGGAAGG 801

```

RESULT 8

CA245555 656 bp mRNA linear EST 25-SEP-2003
SCBGFL5081A03 g Saccharum officinarum FL5 Saccharum officinarum
LOCUS cDNA clone SCBGFL5081A03 5', mRNA sequence.
DEFINITION
ACCESSION CA245555
VERSION CA245555.1 GI:35324301
KEYWORDS EST.
SOURCE Saccharum officinarum

```

ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum.
1 (bases 1 to 656)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
CONTACT: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 081 row: A column: 03
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1..656
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCBGF5081A03"
/lab_host="DH10B"
/clone_lib="Saccharum officinarum FL5"
/note="Organ: Developed inflorescence (20cm-long) without
rachis; Vector: pSport1; Site 1: SalI; Site 2: NotI; An
unidirectional cDNA library generated from [Developed
inflorescence (20cm-long) without rachis]. cDNA was
prepared from polyA+ mRNA using SuperScript Plasmid
System Kit (Invitrogen). The double-strand cDNAs were
fractionated in a sepharose CL-2B 40cm-columns and
fragments sizing between 0.8 and 1.5 Kb were
directionally cloned into the vector. Details of each
source of RNA and library construction can be obtained at
http://sucest.lad.ic.unicamp.br/public"

ORIGIN
Alignment Scores:
Pred. No.: 1,91e-90 Length: 656
Score: 1034.00 Matches: 196
Percent Similarity: 97.60% Conservative: 7
Best Local Similarity: 94.23% Mismatches: 4
Query Match: 67.98% Indels: 1
DB: 14 Gaps: 0

US-10-042-894A-8 (1-289) x CA245555 (1-656)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerLys 20
Db 33 ATGTCGACCTCGCCCGCGGAGCACCAGTCGCGGCCACCGCGCTCCGACACAG 92
QY 21 LeuGlyProLeuLeuAspGlySerGlyLeuPheTyrIysProLeuGlnAlaGlyAspArg 40
Db 93 CTGGTCCGCTCATCGACGCTCTGGCTCTTCTACAAGCGCTCCAGTCGCGGACCGC 152
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
Db 153 GGGGAGACAGAGTCGCTCTTACGAGCGCTTCCACCCACCGCCGCGCGCGCGCG 212
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 213 ATCGGGACACCTTCTTCGCGCGGTTCAGGACGCGACTCTCCACCGCGCGCGCG 272
QY 81 ProGlyGluProHisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaPro 100
Db 273 CCAGGGAGCGCATCGCACCTCGTCTGACGACCTCTCGCGGGCTGGAGGCGCC 332
QY 101 CysValAlaAspIleLysIleGlyAlaIleThrProSerSerProGluProTyr 120

```

```

Db 333 TGCGTCCGCGACATCAAGATCGCGGCATCACTGCGCGCGCGAGCTCGCGAGCCCTAC 392
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 393 GTGCCCAAGTGCCTCGCAAGGACCGCGGACCAACGAGCATTCTGCTCGGATTCGCGGTC 452
QY 141 SerGlyValArgValValGlyProGluGlyValAlaValTrpArgThrGluArgProGluVal 160
Db 453 TCCGGGTCGCGGTCTGTCGTCGCCGAGGCGCGTGTGGCGGACTGAGCGCCCGGAGGTG 512
QY 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSerValAla 180
Db 513 AAGGCGCTGGACACCGCGCGTCCGCGCGCTCTCGCGCGCTACGTTTCATCGGTGCC 572
QY 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGly-LysGlyGlyValLe 200
Db 573 GACGAGGAGATGAGCTGCGCACTCGCGCGCGGTGTACGCGCGCAAAAGGGGAGTCTT 632
QY 200 uSerGlnLeuArgGluLeuLys 207
Db 633 GTCACAGCTGCGGAGCTTAAG 654

RESULT 9
CA130685
LOCUS 722 bp mRNA linear EST 24-SEP-2003
DEFINITION SCCRT1004G05.g RT1 Saccharum officinarum cDNA clone SCCRT1004G05
5', mRNA sequence.
ACCESSION CA130685.1 GI:35013914
VERSION CA130685
KEYWORDS Saccharum officinarum
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
REFERENCE Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
AUTHORS The libraries that made SUCEST
TITLE Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
JOURNAL Contact: Arruda P
COMMENT Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 004 row: G column: 05
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1..722
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCCRT1004G05"
/lab_host="DH10B"
/clone_lib="RT1"
/note="Organ: Root tips (0.3cm-long) from adult plants;
Vector: pSport1; Site 1: SalI; Site 2: NotI; An
unidirectional cDNA library generated from [Root tips
(0.3cm-long) from adult plants]. cDNA was prepared from
polyA+ mRNA using SuperScript plasmid System Kit
(Invitrogen). The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 Kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
http://sucest.lad.ic.unicamp.br/public"

ORIGIN
Alignment Scores:

```

```

Pred. No.: 1.97e-86 Length: 722
Score: 993.50 Matches: 202
Percent Similarity: 92.07% Conservative: 7
Best Local Similarity: 88.99% Mismatches: 17
Query Match: 65.32% Indels: 4
DB: 13 Gaps: 0

US-10-042-894A-8 (1-289) x CA130685 (1-722)
QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerLys 20
Db 31 ATGTCGACCTCGC-CGCGCGAGCACAAGTCGCGCCACCGCGCTCCGCAACAG 89
QY 21 LeuGlyProLeuLeuAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
Db 90 CTGGTCCGCTCATCGACGCTCTGGCTCTTCTACAAGCGCTCCAGGTCCGGGACCGC 149
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
Db 150 GGGAGACGAGCTCGGCTTCTACGAGCGCTTCTCCACCGCGCGCTCCGCGCCGC 209
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 210 ATCGGGACACCTTCTCCCGCGGTTTCACGGACGGATCTCTCCACCGAGCGCG 269
QY 81 ProGlyGluProHisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaPro 100
Db 270 CCAGGGAGCGCATCGCACCTCGTCTCGACGACCTCTCGCGGGCTGGAGGCGCC 329
QY 101 CysValAlaAspLeuLysIleGlyAlaIleThrTrpProSerProGluProTyr 120
Db 330 TGGTCCGCGACATCAAGATCGGCGCATCATCGTGGCGCGGAGCTCGCGGAGCCCTAC 389
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 390 GTGCCCAAGTGCCTCGCCAGGACCGGGACACGAGCATCTGCTCGGATTCGGGTC 449
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
Db 450 TCCGGCTCCGGTCTCGTCCGTCGAGGGCGCGTGTGGCGGACTGAGCGCGCGAGGTG 509
QY 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSerValAla 180
Db 510 AAGCCCTGTGACACCGCGCGCTCGCGCGTGTCTCGCGCGCTTACGTTATCCGTTGCC 569
QY 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTrpGlyGlyGlyValle 200
Db 570 GACGAGGGATGGACTCGGCACTCGCGCGCGGGGACCGCGGCAAGGGGAGTCTT 629
QY 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220
Db 630 GTCAAGCTGCCGAGCTTAAAGCGTGAAGCGTGGTTCGAGGA-CAAACTCTGGATCACTT-TACTC 687
QY 220 rAlaSerIleLeuLeuGly 226
Db 688 GGGCGCATCTCTTTCGGGC 706

RESULT 10
CG776236 591 bp DNA linear GSS 29-OCT-2003
LOCUS 112305E08.x1 1123 - RescueMu Grid L Zea mays genomic, genomic
DEFINITION survey sequence.
ACCESSION CG776236
VERSION CG776236.1 GI:38034108
KEYWORDS GSS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 591)
AUTHORS Walbot,V.
TITLE Maize genomic sequences found using engineered RescueMu transposon
```

JOURNAL
COMMENT

Unpublished (2001)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 1123005 column: 3
Class: transposon-tagged.

FEATURES
source

/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="mixed background W23/A188/B73/K55"
/db_xref="taxon:4577"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="1123 - RescueMu Grid L"
/notes="Organ: leaf; Vector: RescueMu (engineered from
pBluescript backbone); Site 1: BamHI; Site 2: BglII;
RescueMu is a 4.9 kb, modified maize Mu transposon
designed to allow plasmid rescue from total genomic DNA.
Mu elements insert preferentially into transcription
units. For more information on RescueMu, go to the web
site 'www.zmdb.iastate.edu' and follow the links for
'RescueMu.' Grid L was grown in Molokai in 2001. DNA was
extracted from leaf strips, double digested using BamHI
and BglII, and ligated to form circular plasmids. DH10B
cells were transformed and then screened on LB plates with
ampicillin."

ORIGIN

Alignment Scores:
Pred. No.: 2.03e-86 Length: 591
Score: 992.00 Matches: 187
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.47% Mismatches: 0
Query Match: 65.22% Indels: 0
DB: 29 Gaps: 0

US-10-042-894A-8 (1-289) x CG776236 (1-591)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerLys 20
Db 28 ATGTCGACCTCCACCGCGGAGCACAAGTCGCGCCACCGCGCTCCGCGAGCAAG 87
QY 21 LeuGlyProLeuLeuAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
Db 88 CTGGGCGCGCTCATCGACGCTCCGCGCTTCTTACAAGCGCGCTCCGCGCGGACCGCT 147
QY 41 GlyGluHisGlnValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
Db 148 GGGAGACGAGGTGCGCTTCTATGAGCGTCTTCGCGCCACCGCGCGCTCCGCGCCGC 207
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 208 ATCCGAGACACCTTCTTCCCGCGGTTCACGCGACGCGACTCTCTCCCAACGAGGCGAG 267
QY 81 ProGlyGluProHisProHisLeuValLeuAspLeuLeuAlaGlyPheGlnAlaPro 100
Db 268 CCCGGGAGCGCATCTCTACCTCGCTCGAGACCTCTCTCGGGGTTTCAAGGCGCC 327
QY 101 CysValAlaAspIleLysIleGlyAlaIleThrTrpProSerProGluProTyr 120
Db 328 TGGTCCGACATCAAGATCGGCGCATCATCGTGGCCACCGAGTTCGCGGAGCCCTAC 387
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 388 ATCGCAAGTGCCTCGCCAAAGGACCGCGGACCAACGAGCGCTTCTGCTCGGATTCGCGGTC 447
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160

Db	448	TCGGGGTTCGAGTCGTCGGCCCGGAGCGCGGTGGCGGACGAGCGCCCGGAGTG	507
QY	161	LysAlaMetAspThrAlaGlyValArgValLeuArgTyrValSerSerValAla	180
Db	508	AAGGCATGACACCGCGCGTCCGCGGTGCTCGGCGCTACGTGTCATCCGTGCC	567
QY	181	AspGluGlyMetAspCysAlaLeu	188
Db	568	GACGAGGGATGGACTGTGGCTC	591
RESULT 11			
LOCUS	CA202412	671 bp	mRNA linear EST 25-SBP-2003
DEFINITION	SCRLFL1009H02.9 Flt1 Saccharum officinarum cDNA clone SCRLFL1009H02		
ACCESSION	CA202412		
VERSION	CA202412.1	GI:35237651	
KEYWORDS	EST.		
SOURCE	Saccharum officinarum		
ORGANISM	Saccharum officinarum		
REFERENCE			
AUTHORS	Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.		
TITLE	The libraries that made SUCST		
JOURNAL	Genet. Mol. Biol. 24 (1-4), 1-7 (2001)		
COMMENT	Contact: Arruda P Centro de Biologia Molecular e Engenharia Genetica Universidade Estadual de Campinas Caixa Postal 6010, 13083-970, Campinas SP, Brazil Tel: 55 19 3788 1137 Fax: 55 19 3788 1089 Email: parruda@unicamp.br Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at http://www.bccccenter.fcav.unesp.br Plate: 009 row: H column: 02 Seq primer: T7 Promoter Primer. Location/Qualifiers 1. 671 /organism="Saccharum officinarum" /mol_type="mRNA" /db_xref="taxon:4547" /clone="SCRLFL1009H02" /lab_host="DH10B" /clone_lib="FL1" /note="Organ: Inflorescence at beginning of development (1cm-long); Vector: pSport1; Site_1: SalI; Site_2: NotI; An unidirectional cDNA library generated from [inflorescence at beginning of development (1cm-long)]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sapharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at http://sucst.iad.ic.unicamp.br/public "		
ORIGIN			
Alignment Scores:			
Pred. No.:	7.58e-86	Length:	671
Score:	987.00	Matches:	192
Percent Similarity:	91.93%	Conservative:	13
Best Local Similarity:	86.10%	Mismatches:	17
Query Match:	64.89%	Indels:	2
DB:	14	Gaps:	0
US-10-042-894A-8 (1-289) x CA202412 (1-671)			
QY	41	GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg	60

Db	2	GGGAGACGAGCTCGCTTCTACGAGCGGTTCTCCAAACACACGCGCGTCCCGGCGCGC	61
QY	61	IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln	80
Db	62	ATCCGGGACACCTTCTTCCCGGGTTCACGGACGCGACTCTCTCCACACGAGGCGCG	121
QY	81	ProGlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro	100
Db	122	CCAGGGAGCGGCATCCGCACCTCGTCTCGACGACTCTCTCGCGGGCTGGAGGGCGCC	181
QY	101	CysValAlaAspIleIleGlyAlaIleThrTrpProSerSerProGluProTyr	120
Db	182	TGCTCGCGCATCAAGATCGCGCCATCATCGTGGCCCGGAGCTCGCGGAGCCCTAC	241
QY	121	IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal	140
Db	242	GTCGCCAAGTGCTCGCCAAAGCACCGCGGACCATCTGCTCGGATTCGGATTCGGGTC	301
QY	141	SerGlyValArgValValGlyProGluGlyAlaValTyrArgThrGluArgProGluVal	160
Db	302	TCCGGCGTCCGGTCTCGTCCGAGGCGCGTGTGGCGGACTGAGCGCCGAGGTTG	361
QY	161	LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSerValAla	180
Db	362	AAGGCCCTGGACACCGCGCGTCCGCGCGTCTCCGCGCTAGCTTTCATCCGTTGCC	421
QY	181	AspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyValLeu	200
Db	422	GACGAGGATGAGTGTGCGACTCCCGCGCGGTGTACGGCGGCGGAGGAGTCTTG	481
QY	201	SerGlnLeuArgGluLeuLysAlaTyrPheGluGluGlnThrLeuPheHisPheTyrSer	220
Db	482	TCACAGCTCGCGGAGCTTAAGCGTGGTTCGAGGAGCAGACTCTGCTCCACTTCTACTCG	541
QY	221	AlaSerIleLeuLeuGlyTyrAspAlaAlaAlaValAlaAlaGlyGlyGlyGlyGly	240
Db	542	GCGGTGAATCTTTTGGGCTATGATGCTAGTCAAGTGCAGCAAGCGGAGGTGAAGTGGG	601
QY	241	ValThrValLysLeuValAspPheAlaHisValAlaGluGlyAsp-GlyValIleAspHi	260
Db	602	TTGAAGGTGAA-CTTGTTGACTTTTCCCAATTTTCCCAAGTTGATAGGGGGAATGACCA	660
QY	260	sAsnPhe	262
Db	661	AAATTC	667
RESULT 12			
LOCUS	AY109355	3374 bp	mRNA linear HTC 17-OCT-2002
DEFINITION	Zea mays CL49_2 mRNA sequence.		
ACCESSION	AY109355		
VERSION	AY109355.1	GI:21213015	
KEYWORDS	HTC.		
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE			
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
TITLE	1 (bases 1 to 3374)		
JOURNAL	Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V. Design of Maize Mapping Project/Dupont Consensus Sequences for		
AUTHORS	Overgo Probes		
TITLE	Unpublished (2002)		
JOURNAL	2 (bases 1 to 3374)		
AUTHORS	Coe, E.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA		
COMMENT	If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org ; ZmDB, www.zmdb.iastate.edu ; TIGR, www.tigr.org ; or NCBI, www.ncbi.nlm.nih.gov . When the source of the		

maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES

source
 1. .3374
 /organism="Zea mays"
 /mol_type="mRNA"
 /db_xref="MaizeDB:632179"
 /clone_lib="Maize Mapping Project/DuPont Cornsensus Library"
 /note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

ORIGIN

Alignment Scores:
 Pred. No.: 3.45e-84 Length: 3374
 Score: 981.00 Matches: 205
 Percent Similarity: 71.88% Conservative: 2
 Best Local Similarity: 71.18% Mismatches: 62
 Query Match: 64.50% Indels: 20
 DB: 11 Gaps: 1

US-10-042-894A-8 (1-289) x AY109355 (1-3374)

QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20
 DB 72 ATGCCCGACTCCACCCGCGGAGCACCAAGTCGCGGTACCGCGCTCGCCAGCAAG 131
 QY 21 LeuGlyProLeuLeuAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
 DB 132 CTGGGCCCACTCATCGCAGACTCTGGCTCTCTACAAAGCGCTCCAGCGCGGCGGACCGT 191
 QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
 DB 192 GGGGAGCAGCAGGTCGCTTCTATGAGCGNNNNNNNNNNNNNNNNNNNNNNNNNNNN 251
 QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
 DB 252 NTCCGAGACACTTCTNNNNNGTTCCACGGCAGCGACTCCTCCACCGAGGCGCAG 311
 QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuAlaGlyPheGlnAlaPro 100
 DB 312 CCCGGGAGCGCATCCGCACTCGTCTCTGACGACCTCTCTCGCGGGGTTTGAGCGGCC 371
 QY 101 CysValAlaAspLeuLysIleGlyAlaIleThrTrpProProSerSerProGluProTyr 120
 DB 372 TGCCTCGCAGACATCAAGATCGGTGCCATCACGTG----- 406
 QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
 DB 407 -----ACCACGAGCGTTCGTCTCGGATTCGCGTCC 436
 QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
 DB 437 TCCGGCTCCGAGTCGTCCGCGCGGCGCGCGTNNNNNNNNNNNNNNNNNNNNNNNN 496
 QY 161 LysAlaMetAspThrAlaGlyValArgArgValLeuArgArgTyrValSerSerValAl 180
 DB 497 AAGGCTATGGACATTGTTCGGCGTCCGCGCGTGTCTCGCGCGCTACGTGTATCCGCTGC 556
 QY 180 aAspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyGlyLysGlyGlyValLe 200
 DB 557 CGACGAGGGATGACTCCGCGCTCCGCGCGCGGTGTACGAGAGAAAGTGAGTCTT 616
 QY 200 uSerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSe 220
 DB 617 GTCACAGCTCGCGAGCTCAAGCGTGTGTTTCGAGGGGCAGACTNNNNNNNNNNNNNN 676

QY 220 rAlaSerIleLeuLeuGlyTyrAspAlaAlaValAlaAlaGlyGlyGlyGlyGlyGly 240
 DB 677 NNN 736
 QY 240 yValThrValLysLeuValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHi 260
 DB 737 NNN 796
 QY 260 sAenPheLeuGlyGlyLeuCysSerLeuLeuLysPheValSerAspIleValProGluTh 280
 DB 797 CAACTTCCTCGCGCGGCTCTGCTAGCTGATCAAGTTTGTTCATGACATTTTCCAGAGAC 856
 QY 280 rProHisThrGlnProLeuGly 287
 DB 857 TCCTCAGACGACCCCTTTGGN 878

RESULT 13
 CC724950/c
 LOCUS
 DEFINITION
 CC724950 722 bp DNA linear GSS 23-JUN-2003
 OGLAB23TH ZM 0.7_1.5_KB Zea mays genomic clone ZMMBMA0307C22,
 genomic survey sequence.
 ACCESSION
 CC724950
 VERSION
 CC724950.1 GI:32143883
 KEYWORDS
 GSS.
 SOURCE
 Zea mays
 ORGANISM
 Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 722)
 AUTHORS
 Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
 Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
 Citek,R.W., Nummer,A., Robbins,D. and Lakey,N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Other GSSs: OGLAB23TV
 Contact: Cathy Whitelaw
 TIGR Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TR
 Class: sheared ends.
 FEATURES
 source
 1. .722
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone_lib="ZMMBMA0307C22"
 /clone_lib="ZM 0.7_1.5_KB"
 /notes="Vector: pBSK-; Site 1: HincII; 0.7-1.5 kb
 methylation filtered genomic DNA library"

ORIGIN
 Alignment Scores:
 Pred. No.: 1e-84 Length: 722
 Score: 976.00 Matches: 197
 Percent Similarity: 88.44% Conservative: 2
 Best Local Similarity: 87.56% Mismatches: 7
 Query Match: 64.17% Indels: 20
 DB: 29 Gaps: 1

US-10-042-894A-8 (1-289) x CC724950 (1-722)

QY 66 pheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGlnProGlyGluProHis 85
 DB 721 TTCCCCCGGTTCACGGCAGCGACTCTCCCCACGAGCGCGCCCGGGGCGCGCAT 662
 QY 86 ProHisLeuValLeuAspLeuAlaGlyPheGlnAlaProCysValAlaAspIle 105
 DB 661 CCGCACCTGCTCTCGACGACCTCTCGCGGGGTTTGAGGCGCCCTGCGTCCGAGACATC 602

```
QY 106 LysIleGlyAlaIleThrTrpProSerSerProGluProTyrIleAlaLysCysLeu 125
Db 601 AAGATCGGTGCCATCACGTG-----582
QY 126 AlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgValSerGlyValArgVal 145
Db 581 -----ACACAGCGCTTCGCTCGGATTCGCGTTCGCGGCTCCGAGTC 537
QY 146 ValGlyProGluGlyAlaValTrpArgThrGluArgProGluValLysAlaMetAspThr 165
Db 536 GTCGCGCCCGAGGCGCGCTGTGCGGACGGAGCGCCGAGGTGAGGCTATGACATT 477
QY 166 AlaGlyValArgValLeuArgGlyTyrValSerSer-ValAlaAspGluGlyMetAs 185
Db 476 GTCGCGCTCCGCGCGTGTCTCGCGCTACGTGTATCGCTTGCAGACGAGGGATGGA 417
QY 185 pCysAlaLeuAlaAlaValTyrGlyLysGlyGlyValLeuSerGlnLeuArgG 205
Db 416 CTGCGCGCTCGCGCGCGCGGTGTACGAGGAAAGTGGAGTCTGTACAGTCGCGCA 357
QY 205 uLeuLysAlaIleThrPheGluGlnThrLeuPheHisPheTyrSerAlaSerIleLeu 225
Db 356 GCTCAGGCGTGTTCGAGGGCGAGCTCTGTCCACTTCTACTCGGCTCGATTCTTCT 297
QY 225 uGlyTyrAspAlaAlaValAlaAlaGlyGlyAspGlyGlyValThrValLysLe 245
Db 296 GGGCTATGATGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 237
QY 245 uValAspPheAlaHisValAlaGluGlyAspGlyValIleAspHisAsnPheLeuGly 265
Db 236 GTTGACTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 177
QY 265 yLeuCysSerLeuIleLysPheValSerAspIleValProGluThrProHisThrGlnPr 285
Db 176 GCTCTGTAGCTGATCAAGTTTGTCTGACATTGTTCAGAGTGTTCAGAGTCTCTC 117
QY 285 oLeuGlyProSer 289
Db 116 TTGGGTCTCTCT 104

RESULT 14
CAL130686
LOCUS
DEFINITION CAL130686 744 bp mRNA linear EST 24-SEP-2003
5', mRNA sequence.
ACCESSION CAL130686
VERSION CAL130686.1 GI:35013916
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum.
1 (bases 1 to 744)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bccccenter.fcav.unesp.br
Plate: 004 row: G column: 06
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1..744
/organism="Saccharum officinarum"
```

```
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCCRT1004G06"
/lab_host="DH10B"
/clone_lib="Rtl"
/notes="Morgan: Root tips (0.3cm-long) from adult plants;
Vector: pSPori1; Site_1: SalI; Site_2: NotI; An
unidirectional cDNA library generated from [Root tips
(0.3cm-long) from adult plants]. cDNA was prepared from
polyA+ mRNA using Superscript Plasmid System Kit
[Invitrogen]. The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 Kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
http://sucst.lad.ic.unicamp.br/public"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 2,06e-80 Length: 744
Score: 932.00 Matches: 191
Percent Similarity: 86.02% Conservative: 12
Best Local Similarity: 80.93% Mismatches: 26
Query Match: 61.28% Indels: 9
DB: 13 Gaps: 1

US-10-042-894A-8 (1-289) x CAL130686 (1-744)
QY 1 MetSerAspLeuHisProGluHisGlnValAlaGlyHisArgAlaSerAlaSerLys 20
Db 31 ATGTCGGACCTTCG-CCGCGCGAGACCAAGTCGCGGCCACCGCGCTCCGCCAACAG 89
QY 21 LeuGlyProLeuIleAspGlySerGlyLeuPheTyrLysProLeuGlnAlaGlyAspArg 40
Db 90 CTGGGTCCGCTCATCGACGGCTCTGCTCTTCTACAAAGCCGCTCAGGTCGGGACCGC 149
QY 41 GlyGluHisGluValAlaPheTyrGluAlaPheSerAlaHisAlaValProAlaArg 60
Db 150 GGGGAGACGAGCTCGCTTCTACAGGCGTTCCTCACCCAGCGCGCTCCGCGCCGCG 209
QY 61 IleArgAspThrPhePheProArgPheHisGlyThrArgLeuLeuProThrGluAlaGln 80
Db 210 ATCCGGGACACCTTCTTCCCGCGGTTTCAGGCGACGACTCTCTCCACCGAGCGCGG 269
QY 81 ProGlyGluProHisProHisLeuValLeuAspAspLeuLeuAlaGlyPheGlnAlaPro 100
Db 270 CCAGGGGAGCGCATCCGCACCTCGCTCTCGACGACCTCTCTCGCGGGGCTGAGGCGGCC 329
QY 101 CysValAlaAspIleLysIleGlyValAlaIleThrTrpProSerSerProGluProTyr 120
Db 330 TCGGTCCCGACATCAAGATCGCGCCATCATCGTGGCGCGCGAGCTCGCGGAGGCGCTAC 389
QY 121 IleAlaLysCysLeuAlaLysAspArgGlyThrThrSerValLeuLeuGlyPheArgVal 140
Db 390 GTCGCCAAGTGTCTGCCAAGGACCGCGGACCCACGAGCATTCGTCTGGATTCCGCGTC 449
QY 141 SerGlyValArgValValGlyProGluGlyAlaValTrpArgThrGluArgProGluVal 160
Db 450 TCCGCGCTCCGCGTGTCTGTCGTCGCGCGCGCGCTGGGGCGGACTGAAACCGCGAGGTG 509
QY 161 LysAlaMetAspThrAlaGlyValArgValLeuArgValLeuArgTyrValSerValAla 180
Db 510 AAGGCCCTGGACACCGCGCGCTCCCGCGGGGCTTCGCGCGCTAGCTTTATTCTGTGTC 569
QY 181 AspGluGlyMetAspCysAlaLeuAlaAlaValTyrGlyLysGlyGlyValLeu 200
Db 570 GACGAGGGGATGACTGTCTCCCTCCCGCGCGCGGT-TACGGGGGCAAAAGGGGAAATTG 628
QY 201 SerGlnLeuArgGluLeuLysAlaTrpPheGluGluGlnThrLeuPheHisPheTyrSer 220
Db 629 GTAACATGGGCAAGCTTAAAGGGGGGTAGAGGGGAAAAATTTGGTTC-----676
QY 221 AlaSerIleLeu-LeuGlyTyrAspAlaAlaValAlaAlaGly 235
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2004, 20:51:24 ; Search time 60 Seconds
(without alignments)
1360.937 Million cell updates/sec

Title: US-10-042-894A-8

Perfect score: 1521

Sequence: 1 MSLLHPPEHQVAGHRASAK.....IKFVSIVPETHQTPLGPS 289

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 29Jan04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003ba:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1521	100.0	289	5	Aae26196 Maize ino
2	1485	97.6	289	5	Aae26195 Maize ino
3	914.5	60.1	227	5	Aae26194 Maize ino
4	748.5	49.2	240	5	Aae26193 Maize ino
5	742.5	48.8	295	5	Aae26199 P. argent
6	722.5	47.5	310	5	Aae26198 Eucalyptu
7	714.5	47.0	300	3	Aag46591 Arabidops
8	714.5	47.0	330	3	Aag46590 Arabidops
9	700.5	46.1	300	3	Aag44257 Arabidops
10	700.5	46.1	330	3	Aag44256 Arabidops
11	674.5	44.3	286	3	Aag19913 Arabidops
12	674.5	44.3	286	3	Aag20765 Arabidops
13	674.5	44.3	286	3	Aag06976 Arabidops
14	674.5	44.3	327	3	Aag06975 Arabidops
15	673.5	44.3	286	5	Abb93390 Herbicida
16	644.5	42.4	279	5	Aae26197 Soybean i
17	578	38.0	111	5	Aae26200 Maize ino
18	577	37.9	111	5	Aae26201 Maize ino
19	542.5	35.7	248	3	Aag46592 Arabidops
20	494	32.5	205	3	Aag19915 Arabidops
21	494	32.5	205	3	Aag20767 Arabidops
22	494	32.5	206	3	Aag20766 Arabidops
23	494	32.5	206	3	Aag19914 Arabidops
24	494	32.5	206	3	Aag06977 Arabidops
25	439.5	28.9	202	3	Aag24258 Arabidops

26	265	17.4	231	4	Abb68470 Drosophil
27	201.5	13.2	416	6	Aae34863 Human kin
28	192	12.6	410	5	Aae24144 Human ino
29	192	12.6	410	7	Adc06743 Human pro
30	191	12.6	293	6	Ada55092 Human pro
31	183	12.0	353	4	Aam41451 Human pol
32	183	12.0	426	4	Aam39665 Human pol
33	183	12.0	426	4	Abb94867 Human pro
34	183	12.0	426	4	Aag89197 Human sec
35	183	12.0	426	4	Aag89190 Human sec
36	183	12.0	426	7	Adc06747 Human ino
37	183	12.0	481	4	Aag89295 Human sec
38	182.5	12.0	396	7	Adc06749 Marine in
39	177	11.6	426	3	Abb42986 Human ORF
40	163.5	10.7	441	7	Adc06745 Human ino
41	147	9.7	477	5	Abb09782 Amino aci
42	145	9.5	461	5	Abb09772 Amino aci
43	142.5	9.4	441	5	Abb09770 Amino aci
44	142.5	9.4	441	5	Abb09780 Amino aci
45	139	9.1	313	4	Aao07085 Human pol

ALIGNMENTS

RESULT 1
AAE26196
ID AAE26196 standard; protein; 289 AA.
XX
AC AAE26196;
XX
DT 14-NOV-2002 (first entry)
XX
DE Maize inositol polyphosphate kinase (IPPK) #4.
XX
KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
KW nutritional value; animal feed; transgenic.
XX
OS Zea mays.
XX
PN WO200259324-A2.
XX
PD 01-AUG-2002.
XX
PF 09-JAN-2002; 2002WO-US003120.
XX
PR 12-JAN-2001; 2001US-0261465P.
PR (PION-) PIONEER HI-BRED INT INC.
XX
PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX
DR WPI; 2002-636540/68.
XX
N-PSDB; AAD43514.
PT New inositol polyphosphate kinase polynucleotides and polypeptides,
PT useful in modulating phytic acid biosynthesis by decreasing phytate or
PT increasing non-phytate phosphorous to improve the nutritional value of
PT animal feed.
PS Claim 12; Page 65-66; 86pp; English.
XX
CC The invention relates to novel inositol polyphosphate kinase (IPPK)
CC polypeptides and polynucleotides. Sequences of the invention are useful
CC in modulating the phytic acid biosynthesis by decreasing phytate and/or
CC increasing non-phytate phosphorous to improve the nutritional value of
CC animal feed, or to reduce the environmental impact of animal waste.
CC Polynucleotides of the invention are to produce transgenic plants with an
CC altered phenotype. IPPK proteins are used to screen compounds that
CC modulate their activity and raising anti-idiotypic antibodies. The
CC present sequence is maize IPPK protein
XX
SQ Sequence 289 AA;

Query Match 100.0%; Score 1521; DB 5; Length 289;
 Best Local Similarity 100.0%; Pred. No. 5.2e-159;
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLLHPEHGVAGHRASAKGLPLIDSGLFYKPLQAGDRGEHEVAFYFAFSAHAAPAR 60
 |||||
 Db 1 MSLLHPEHGVAGHRASAKGLPLIDSGLFYKPLQAGDRGEHEVAFYFAFSAHAAPAR 60
 |||||

QY 61 IRDTFFPRFHGTRLLPTEAQPGEPHPLVLDLLAGFQAPCVADIKIGAITWPPSSPEPY 120
 |||||
 Db 61 IRDTFFPRFHGTRLLPTEAQPGEPHPLVLDLLAGFQAPCVADIKIGAITWPPSSPEPY 120
 |||||

QY 121 IAKCLAKDRGTTSVLLGFRVSGVRVVGPEGAVWRTERPEVKAMDTAGVRRVLRYYSSVA 180
 |||||
 Db 121 IAKCLAKDRGTTSVLLGFRVSGVRVVGPEGAVWRTERPEVKAMDTAGVRRVLRYYSSVA 180
 |||||

QY 181 DEGMDCALAAAVYGGKGVLSQRELKAWFEQTLFHFYSASILLGYDAAAAGGDDGGG 240
 |||||
 Db 181 DEGMDCALAAAVYGGKGVLSQRELKAWFEQTLFHFYSASILLGYDAAAAGGDDGGG 240
 |||||

QY 241 VTKLVDFAHVAGDGVIDHNFGLGGLCSLIKFSVDIVPETHPTQPLGPS 289
 |||||
 Db 241 VTKLVDFAHVAGDGVIDHNFGLGGLCSLIKFSVDIVPETHPTQPLGPS 289
 |||||

RESULT 2
 AAE26195
 ID AAE26195 standard; protein; 289 AA.
 XX
 AC AAE26195;
 XX
 DT 14-NOV-2002 (first entry)
 XX
 DE Maize inositol polyphosphate kinase (IPPK) #3.
 XX
 KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
 KW nutritional value; animal feed; transgenic.
 XX
 OS Zea mays.
 XX
 PN WO200259324-A2.
 XX
 PD 01-AUG-2002.
 XX
 PF 09-JAN-2002; 2002WO-US003120.
 XX
 PR 12-JAN-2001; 2001US-0261465P.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
 XX
 DR WPI; 2002-636540/68.
 DR N-PSDB; AAD43513.
 XX
 PT New inositol polyphosphate kinase polynucleotides and polypeptides,
 PT useful in modulating phytic acid biosynthesis by decreasing phytate or
 PT increasing non-phytate phosphorous to improve the nutritional value of
 PT animal feed.
 XX
 PS Claim 12; Page 63-64; 86pp; English.
 XX
 CC The invention relates to novel inositol polyphosphate kinase (IPPK)
 CC polypeptides and polynucleotides. Sequences of the invention are useful
 CC in modulating the phytic acid biosynthesis by decreasing phytate and/or
 CC increasing non-phytate phosphorous to improve the nutritional value of
 CC animal feed, or to reduce the environmental impact of animal waste.
 CC Polynucleotides of the invention are to produce transgenic plants with an
 CC altered phenotype. IPPK proteins are used to screen compounds that
 CC modulate their activity and raising anti-idiotypic antibodies. The
 CC present sequence is maize IPPK protein

SQ Sequence 289 AA;
 Query Match 97.6%; Score 1485; DB 5; Length 289;
 Best Local Similarity 97.9%; Pred. No. 5e-155;
 Matches 283; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSLLHPEHGVAGHRASAKGLPLIDSGLFYKPLQAGDRGEHEVAFYFAFSAHAAPAR 60
 |||||
 Db 1 MPDLHPEHGVAGHRASAKGLPLIDSGLFYKPLQAGDRGEHEVAFYFAFSAHAAPAR 60
 |||||

QY 61 IRDTFFPRFHGTRLLPTEAQPGEPHPLVLDLLAGFQAPCVADIKIGAITWPPSSPEPY 120
 |||||
 Db 61 IRDTFFPRFHGTRLLPTEAQPGEPHPLVLDLLAGFQAPCVADIKIGAITWPPSSPEPY 120
 |||||

QY 121 IAKCLAKDRGTTSVLLGFRVSGVRVVGPEGAVWRTERPEVKAMDTAGVRRVLRYYSSVA 180
 |||||
 Db 121 IAKCLAKDRGTTSVLLGFRVSGVRVVGPEGAVWRTERPEVKAMDTAGVRRVLRYYSSVA 180
 |||||

QY 181 DEGMDCALAAAVYGGKGVLSQRELKAWFEQTLFHFYSASILLGYDAAAAGGDDGGG 240
 |||||
 Db 181 DEGMDCALAAAVYGGKGVLSQRELKAWFEQTLFHFYSASILLGYDAAAAGGDDGGG 240
 |||||

QY 241 VTKLVDFAHVAGDGVIDHNFGLGGLCSLIKFSVDIVPETHPTQPLGPS 289
 |||||
 Db 241 VTKLVDFAHVAGDGVIDHNFGLGGLCSLIKFSVDIVPETHPTQPLGPS 289
 |||||

RESULT 3
 AAE26194
 ID AAE26194 standard; protein; 227 AA.
 XX
 AC AAE26194;
 XX
 DT 14-NOV-2002 (first entry)
 XX
 DE Maize inositol polyphosphate kinase (IPPK) #2.
 XX
 KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
 KW nutritional value; animal feed; transgenic.
 XX
 OS Zea mays.
 XX
 PN WO200259324-A2.
 XX
 PD 01-AUG-2002.
 XX
 PF 09-JAN-2002; 2002WO-US003120.
 XX
 PR 12-JAN-2001; 2001US-0261465P.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
 XX
 DR WPI; 2002-636540/68.
 DR N-PSDB; AAD43512.
 XX
 PT New inositol polyphosphate kinase polynucleotides and polypeptides,
 PT useful in modulating phytic acid biosynthesis by decreasing phytate or
 PT increasing non-phytate phosphorous to improve the nutritional value of
 PT animal feed.
 XX
 PS Claim 12; Page 61-62; 86pp; English.
 XX
 CC The invention relates to novel inositol polyphosphate kinase (IPPK)
 CC polypeptides and polynucleotides. Sequences of the invention are useful
 CC in modulating the phytic acid biosynthesis by decreasing phytate and/or
 CC increasing non-phytate phosphorous to improve the nutritional value of
 CC animal feed, or to reduce the environmental impact of animal waste.
 CC Polynucleotides of the invention are to produce transgenic plants with an
 CC altered phenotype. IPPK proteins are used to screen compounds that
 CC modulate their activity and raising anti-idiotypic antibodies. The
 CC present sequence is maize IPPK protein

```

XX SQ Sequence 227 AA;
Query Match 60.18; Score 914.5; DB 5; Length 227;
Best Local Similarity 78.5%; Pred. No. 3.8e-92;
Matches 183; Conservative 8; Mismatches 35; Indels 7; Gaps 2;

QY 1 MSDLHPPEHQVAGHRASASKGLPLIDGSLFYKPLQAGDRGEHEVAFYFAFSAHAAPAR 60
Db 1 MPDLHPPEHQVAGHRASASKGLPLIDGSLFYKPLQAGDRGEHEVAFYFAFSAHAAPAR 60
QY 61 IRDTFFPRFGTLLPTEAQGPGEHPHLVLDLLAGFQPCVADIKIGAITWPPSSPEPY 120
Db 61 IRDTFFPRFGTLLPTEAQGPGEHPHLVLDLLAGFEAPCVADIKIGAITWPPSSPEPY 120
QY 121 IAKCLAKDRGTTSVLLGFRVVGPEGAVWTERPEVKAMDTAGVRVLRVYSSVA 180
Db 121 IAKCLAKDRGTTSVLLGFRVVGPEGAVWTERPEVKAMDTAGVRVLRVYSSVA 180
QY 181 DEGMDCALAAAVYGGKGVLSQLRELKAWFEEQTLPH-FYSASILLGYDAAAV 232
Db 180 -----CRGDLRARGGVGRKWSLVTRAAQGVVGRGAASVPLLIGVDSGL 227

RESULT 4
AAE26193
ID AAE26193 standard; protein; 240 AA.
AC AAE26193;
XX
XX
DT 14-NOV-2002 (first entry)
XX
DE Maize inositol polyphosphate kinase (IPPK) #1.
XX
KW Maize; enzyme; inositol polyphosphate kinase; IPPK; phytic acid;
KW nutritional value; animal feed; transgenic.
XX
XX Zea mays.
XX
XX WO200259324-A2.
XX
XX 01-AUG-2002.
XX
XX 09-JAN-2002; 2002WO-US003120.
XX
XX 12-JAN-2001; 2001US-0261465P.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX WPI; 2002-636540/68.
XX N-PSDB; AAD43511.
XX
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
XX useful in modulating phytic acid biosynthesis by decreasing phytate or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed.
XX
XX Claim 12; Page 59-60; 86pp; English.
XX
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
XX polypeptides and polynucleotides. Sequences of the invention are useful
XX in modulating the phytic acid biosynthesis by decreasing phytate and/or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed, or to reduce the environmental impact of animal waste.
XX Polynucleotides of the invention are to produce transgenic plants with an
XX altered phenotype. IPPK proteins are used to screen compounds that
XX modulate their activity and raising anti-idiotypic antibodies. The
XX present sequence is maize IPPK protein
XX
XX Sequence 240 AA;

Query Match 49.2%; Score 748.5; DB 5; Length 240;
Best Local Similarity 64.7%; Pred. No. 9.1e-74;
Matches 163; Conservative 10; Mismatches 50; Indels 29; Gaps 6;

QY 1 MSDLHPPEHQVAGHRASASKGLPLIDGSLFYKPLQAGDRGEHEVAFYFAFSAHAAPAR 60
Db 1 MPDLHPPEHQVAGHRASASKGLPLIDGSLFYKPLQAGDRGEHEVAFYFAFSAHAAPAR 60
QY 61 IRDTFFPRFGTLLPTEAQGPGEHPHLVLDLLAGFQPCVADIKIGAITWPPSSPEPY 120
Db 61 IRDTFFPRFGTLLPTEAQGPGEHPHLVLDLLAGFEAPCVADIKIGAITWPPSSPEPY 120
QY 121 IAKCLAKDRGTTSVLLGFRVVGPEGAVWTERPEVKAMDTAGVRVLRVYSSVA 164
Db 121 IAKCLAKDRGTTSVLLGFRVLRPSRRPRGERVADGAPGGEGYGHRRPPAPALRVIRLP 180
QY 165 TAG-VRRVLRVYSSVADGMDCALAAAVYGGKGVLSQLRELKAWFEEQTLPHFYSASI 223
Db 181 TRGWTARSRRCCTEEKVESCHSC--ASSRHGWSRLCS-----TSTRRRFFWAMM 228

224 LLGYDAAVAAAG 235
229 LLQSQQAQEVGVG 240

RESULT 5
AAE26199
ID AAE26199 standard; protein; 295 AA.
XX
XX AC AAE26199;
XX
XX 14-NOV-2002 (first entry)
XX
XX P. argentatum inositol polyphosphate kinase (IPPK) #1.
XX
XX Enzyme; inositol polyphosphate kinase; IPPK; phytic acid; transgenic;
XX nutritional value; animal feed.
XX
XX Parthenium argentatum.
XX
XX WO200259324-A2.
XX
XX 01-AUG-2002.
XX
XX 09-JAN-2002; 2002WO-US003120.
XX
XX 12-JAN-2001; 2001US-0261465P.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX WPI; 2002-636540/68.
XX N-PSDB; AAD43517.
XX
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
XX useful in modulating phytic acid biosynthesis by decreasing phytate or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed.
XX
XX Claim 12; Page 72-73; 86pp; English.
XX
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
XX polypeptides and polynucleotides. Sequences of the invention are useful
XX in modulating the phytic acid biosynthesis by decreasing phytate and/or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed, or to reduce the environmental impact of animal waste.
XX Polynucleotides of the invention are to produce transgenic plants with an
XX altered phenotype. IPPK proteins are used to screen compounds that
XX modulate their activity and raising anti-idiotypic antibodies. The
XX present sequence is P. argentatum IPPK protein
XX
XX Sequence 295 AA;

```

```

Query Match      48.8%; Score 742.5; DB 5; Length 295;
Best Local Similarity 52.1%; Pred. No. 5.7e-73;
Matches 149; Conservative 45; Mismatches 79; Indels 13; Gaps 5;

QY 4 LHPPEHQVAGHRASAKLGLPLIDSGGLFYKPLQAGDRGEHEVAFYEAFAAHPARIRD 63
Db 2 LKAPDQVAGHEAGLGLPLIDSGGLFYKPLQAGDRGEHEVAFYEAFAAHPARIRD 60
QY 64 TFFPRFHGTRLLPTEAQFGEPHPLVLDLLAGFAQPCVADIKIGAITWPPSSPEPIAK 123
Db 61 KFFPIYVGTIM--KASTGSDPHMWLQDLTSAHVNPVMDIKIGSRTWAPASEAYIAK 118
QY 124 CLAKDRGTTSVLLGFRVSGVRVGPGEA-VWRTERPEVKAMDTAGVRRVLRVYSSVADE 182
Db 119 CLKKDRSTSLPLGFRISGLQVYDDGSGFYKPHRNYMKTGPADVRLLRKFKVSSNPSA 178
QY 183 -----GMDCALAAAVYGGKGVLSQRLKAWFEEQTLPHFYASILLGYDAAVA 234
Db 179 EMWRTGLGDCSLASFYGGPGLIQLQMLKTFWEDQTIYHYACSFLEIFEKRLVLK 238
QY 235 GGDGGVTVKLVDPFAHVAAGDVIDHNFGLGCLSLIKFVSDIVPET 280
Db 239 GA-RSNAEVKLDPAHVTGNGVIDHNFGLGCLSLIKFISDILSET 283

RESULT 6
AAE26198
ID AAE26198 standard; protein; 310 AA.
AC AAE26198;
XX
XX
DT 14-NOV-2002 (first entry)
DE
DE Eucalyptus grandis inositol polyphosphate kinase (IPPK).
KW Enzyme; inositol polyphosphate kinase; IPPK; phytic acid; transgenic;
KW nutritional value; animal feed.
XX
XX Eucalyptus grandis.
OS
XX
XX WO200259324-A2.
XX
XX 01-AUG-2002.
XX
XX 09-JAN-2002; 2002WO-US003120.
XX
XX 12-JAN-2001; 2001US-0261465P.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Shi J, Beach LR, Wang H, Rafalski JA, Cahoon RE;
XX
XX WPI; 2002-636540/68.
XX
XX N-PSDB; RAD43516.
XX
XX New inositol polyphosphate kinase polynucleotides and polypeptides,
XX useful in modulating phytic acid biosynthesis by decreasing phytate or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed.
XX
XX Claim 12; Page 70-71; 86pp; English.
XX
XX The invention relates to novel inositol polyphosphate kinase (IPPK)
XX polypeptides and polynucleotides. Sequences of the invention are useful
XX in modulating the phytic acid biosynthesis by decreasing phytate and/ or
XX increasing non-phytate phosphorous to improve the nutritional value of
XX animal feed, or to reduce the environmental impact of animal waste.
XX Polynucleotides of the invention are to produce transgenic plants with an
XX altered phenotype. IPPK proteins are used to screen compounds that
XX modulate their activity and raising anti-idiotypic antibodies. The
XX present sequence is Eucalyptus grandis ippk protein

SQ Sequence 310 AA;
Query Match      47.5%; Score 722.5; DB 5; Length 310;
Best Local Similarity 52.8%; Pred. No. 9.9e-71;
Matches 149; Conservative 38; Mismatches 84; Indels 11; Gaps 6;

QY 4 LHPPEHQVAGHRASAKLGLPLIDSGGLFYKPLQAGDRGEHEVAFYEAFAAHPARIRD 63
Db 2 LKVPDQVAGHRGCGKGLGPLVDSDSGRFGYKPLQSDHRGDETEVAFYFYSNTLIPGHIR- 60
QY 64 TFFPRFHGTRLLPTEAQFGEPHPLVLDLLAGFAQPCVADIKIGAITWPPSSPEPIA 122
Db 61 KFFPAFHGTKTI--EASDGSQPQPHLVLEDLVSGRTNPSLMDIKTSGSRTWPEASEEYIQ 118
QY 123 KCLAKDRGTTSVLLGFRVSGVRV-VGEGAVWTERPEVKAMDTAGVRRVLRVYSSVAD 181
Db 119 KCLEKDRNSTSVSLGFRISGLRVYQNSEAGFWQPEKKVYVSFNADGVRSLRKFVSSNLS 178
QY 182 EG---MDCALAAAVYGGKGVLSQRLKAWFEEQTLPHFYASILLGYDAAVAAGGD 237
Db 179 LGENVDPDCLYASKVYCHRGGLIQLQMLKTFWEDQTIYHYACSFLEIFEKRLVLK 236
QY 238 GGGVTVKLVDPFAHVAAGDVIDHNFGLGCLSLIKFVSDIVPE 279
Db 237 CAHPKVKLVDPFAHVMGDGVIDHNFGLGCLSVIKFIRDIAD 278

RESULT 7
AAG46591
ID AAG46591 standard; protein; 300 AA.
AC AAG46591;
XX
XX
DT 18-OCT-2000 (first entry)
DE
DE Arabidopsis thaliana protein fragment SEQ ID NO: 58631.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX
XX 05-MAR-1999; 99US-0123180P.
XX
XX 09-MAR-1999; 99US-0123548P.
XX
XX 23-MAR-1999; 99US-0125788P.
XX
XX 25-MAR-1999; 99US-0126264P.
XX
XX 29-MAR-1999; 99US-0126785P.
XX
XX 01-APR-1999; 99US-0127462P.
XX
XX 06-APR-1999; 99US-0128034P.
XX
XX 08-APR-1999; 99US-0128714P.
XX
XX 16-APR-1999; 99US-0129845P.
XX
XX 19-APR-1999; 99US-0130077P.
XX
XX 21-APR-1999; 99US-0130449P.
XX
XX 23-APR-1999; 99US-0130510P.
XX
XX 23-APR-1999; 99US-0130891P.
XX
XX 28-APR-1999; 99US-0131443P.
XX
XX 30-APR-1999; 99US-0132048P.
XX
XX 04-MAY-1999; 99US-0132407P.
XX
XX 05-MAY-1999; 99US-0132484P.
XX
XX 06-MAY-1999; 99US-0132485P.
XX
XX 06-MAY-1999; 99US-0132486P.
XX
XX 07-MAY-1999; 99US-0132487P.
XX
XX 11-MAY-1999; 99US-0132863P.
XX
XX 14-MAY-1999; 99US-0134213P.
```

PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 28-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 08-JUN-1999; 99US-0138454P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 24-JUN-1999; 99US-0140823P.
PR 28-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.

PR	25-OCT-1999;	99US-0161404P.	PR	30-APR-1999;	99US-0132048P.
PR	25-OCT-1999;	99US-0161405P.	PR	30-APR-1999;	99US-0132407P.
PR	25-OCT-1999;	99US-0161406P.	PR	04-MAY-1999;	99US-0132484P.
PR	26-OCT-1999;	99US-0161359P.	PR	05-MAY-1999;	99US-0132485P.
PR	26-OCT-1999;	99US-0161360P.	PR	06-MAY-1999;	99US-0132486P.
PR	26-OCT-1999;	99US-0161361P.	PR	06-MAY-1999;	99US-0132487P.
PR	28-OCT-1999;	99US-0161920P.	PR	07-MAY-1999;	99US-0132863P.
PR	28-OCT-1999;	99US-0161922P.	PR	11-MAY-1999;	99US-0134256P.
PR	28-OCT-1999;	99US-0161993P.	PR	14-MAY-1999;	99US-0134218P.
PR	29-OCT-1999;	99US-01621142P.	PR	14-MAY-1999;	99US-0134219P.
Query Match 47.0%; Score 714.5; DB 3; Length 300;					
Best Local Similarity 52.8%; Pred. No. 7.2e-70;					
Matches 150; Conservative 39; Mismatches 82; Indels 13; Gaps 8;					
Qy	4	LHPPEHQVAGHRAASKLPLIDGSLFYKPLQAGDRGEHEVAFYFAFSAHAAPARIRD	63	PR	20-MAY-1999;
Db	2	LKVPEHQVAGHIAIDGKGLPLVDQGRFFKPLQGDSRGEHEAFYESFTSNMKVPDHIH-	60	PR	21-MAY-1999;
Qy	64	TFRRFHGTRLTPTEAQGPBH-PHLVLDLLAGFQPCVADIKIGALTWPPSSPPFYIA	122	PR	24-MAY-1999;
Db	61	RYFPVYHGTQIV--EASDGSGLPHLVLDVVGYPANPSVMDVKIGRTWYPDVSEYFK	118	PR	25-MAY-1999;
Qy	123	KCLAKDRGTTSVILGFRVSGVRVVG-PEGAVWRTERPEVKAMDTAGVRRVLRXYVS--SV	179	PR	27-MAY-1999;
Db	119	KCIKKDRQTTVSLGFRVSGFKIFDHQESSFWRAEKKLVLYGNADGARLALRKFSVNSP	178	PR	03-JUN-1999;
Qy	180	ADEGM--DCALAAVYGGKGVLSQLRELKAWFEQTLFHYFYSAILGYDAAV-AAGG	236	PR	08-JUN-1999;
Db	179	ADSNLTNCAPASVYGGCGILAQLELKDWFETQLYHFNSCSILMIYENESILMQGG	238	PR	10-JUN-1999;
Qy	237	DGG---GVTKLVDFAHVAGDGVLDHNFGLGGCSLIKFSVDIV	277	PR	14-JUN-1999;
Db	239	DDAPAPRAQKLVDFAHVLDGNGVIDHNFGLGLCSFIKFIKDIL	282	PR	16-JUN-1999;
RESULT 8					
AAG46590					
ID	AAG46590 standard; protein; 330 AA.				
XX	AC	AAG46590;	PR	18-JUN-1999;	99US-0139457P.
XX	DT	18-OCT-2000 (first entry)	PR	18-JUN-1999;	99US-0139458P.
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 58630.				
XX	KW	Protein identification; signal transduction pathway; metabolic pathway;	PR	18-JUN-1999;	99US-0139460P.
XX	KW	hybridisation assay; genetic mapping; gene expression control; promoter;	PR	18-JUN-1999;	99US-0139461P.
XX	KW	termination sequence.	PR	18-JUN-1999;	99US-0139462P.
XX	OS	Arabidopsis thaliana.	PR	18-JUN-1999;	99US-0139463P.
XX	PN	EP1033405-A2.	PR	18-JUN-1999;	99US-0139750P.
XX	PD	06-SEP-2000.	PR	18-JUN-1999;	99US-0139763P.
XX	PF	25-FEB-2000; 2000EP-00301439.	PR	18-JUN-1999;	99US-0139817P.
XX	PR	25-FEB-1999;	PR	21-JUN-1999;	99US-0139899P.
PR	05-MAR-1999;	99US-0123180P.	PR	22-JUN-1999;	99US-0139899P.
PR	09-MAR-1999;	99US-0123548P.	PR	23-JUN-1999;	99US-0140353P.
PR	23-MAR-1999;	99US-0125788P.	PR	23-JUN-1999;	99US-0140354P.
PR	25-MAR-1999;	99US-0126264P.	PR	24-JUN-1999;	99US-0140695P.
PR	29-MAR-1999;	99US-0126785P.	PR	28-JUN-1999;	99US-0140823P.
PR	01-APR-1999;	99US-0127462P.	PR	29-JUN-1999;	99US-0140991P.
PR	06-APR-1999;	99US-0128234P.	PR	30-JUN-1999;	99US-0141267P.
PR	08-APR-1999;	99US-0128714P.	PR	01-JUL-1999;	99US-0141842P.
PR	16-APR-1999;	99US-0129845P.	PR	01-JUL-1999;	99US-0142154P.
PR	19-APR-1999;	99US-0130077P.	PR	02-JUL-1999;	99US-0142055P.
PR	21-APR-1999;	99US-0130449P.	PR	06-JUL-1999;	99US-0142390P.
PR	23-APR-1999;	99US-0130510P.	PR	08-JUL-1999;	99US-0142803P.
PR	23-APR-1999;	99US-0130891P.	PR	09-JUL-1999;	99US-0142920P.
PR	28-APR-1999;	99US-0131449P.	PR	12-JUL-1999;	99US-0142977P.
PR	28-APR-1999;	99US-01314484P.	PR	13-JUL-1999;	99US-0143542P.
PR	28-APR-1999;	99US-013144814P.	PR	14-JUL-1999;	99US-0143624P.
PR	28-APR-1999;	99US-013144814P.	PR	15-JUL-1999;	99US-0144005P.
PR	28-APR-1999;	99US-013144814P.	PR	16-JUL-1999;	99US-0144085P.
PR	28-APR-1999;	99US-013144814P.	PR	16-JUL-1999;	99US-0144086P.
PR	28-APR-1999;	99US-013144814P.	PR	19-JUL-1999;	99US-0144325P.
PR	28-APR-1999;	99US-013144814P.	PR	19-JUL-1999;	99US-0144331P.
PR	28-APR-1999;	99US-013144814P.	PR	19-JUL-1999;	99US-0144332P.
PR	28-APR-1999;	99US-013144814P.	PR	19-JUL-1999;	99US-0144333P.
PR	28-APR-1999;	99US-013144814P.	PR	19-JUL-1999;	99US-0144334P.
PR	28-APR-1999;	99US-013144814P.	PR	19-JUL-1999;	99US-0144335P.
PR	28-APR-1999;	99US-013144814P.	PR	20-JUL-1999;	99US-0144352P.
PR	28-APR-1999;	99US-013144814P.	PR	20-JUL-1999;	99US-0144632P.
PR	28-APR-1999;	99US-013144814P.	PR	20-JUL-1999;	99US-0144884P.
PR	28-APR-1999;	99US-013144814P.	PR	21-JUL-1999;	99US-0144814P.

PR 21-JUL-1999;	99US-0145086P.
PR 21-JUL-1999;	99US-0145086P.
PR 22-JUL-1999;	99US-0145085P.
PR 22-JUL-1999;	99US-0145087P.
PR 22-JUL-1999;	99US-0145089P.
PR 22-JUL-1999;	99US-0145192P.
PR 23-JUL-1999;	99US-0145145P.
PR 23-JUL-1999;	99US-0145282P.
PR 23-JUL-1999;	99US-0145224P.
PR 26-JUL-1999;	99US-0145276P.
PR 27-JUL-1999;	99US-0145913P.
PR 27-JUL-1999;	99US-0145918P.
PR 28-JUL-1999;	99US-0145918P.
PR 28-JUL-1999;	99US-0145918P.
PR 02-AUG-1999;	99US-0146386P.
PR 02-AUG-1999;	99US-0146388P.
PR 02-AUG-1999;	99US-0146389P.
PR 03-AUG-1999;	99US-0147038P.
PR 04-AUG-1999;	99US-0147204P.
PR 04-AUG-1999;	99US-0147302P.
PR 05-AUG-1999;	99US-0147192P.
PR 05-AUG-1999;	99US-0147260P.
PR 06-AUG-1999;	99US-0147303P.
PR 06-AUG-1999;	99US-0147416P.
PR 08-AUG-1999;	99US-0147493P.
PR 08-AUG-1999;	99US-0147935P.
PR 11-AUG-1999;	99US-0148171P.
PR 11-AUG-1999;	99US-0148319P.
PR 12-AUG-1999;	99US-0148341P.
PR 13-AUG-1999;	99US-0148565P.
PR 13-AUG-1999;	99US-0148684P.
PR 15-AUG-1999;	99US-0149368P.
PR 17-AUG-1999;	99US-0149175P.
PR 18-AUG-1999;	99US-0149426P.
PR 20-AUG-1999;	99US-0149722P.
PR 20-AUG-1999;	99US-0149723P.
PR 20-AUG-1999;	99US-0149929P.
PR 23-AUG-1999;	99US-0149802P.
PR 23-AUG-1999;	99US-0149930P.
PR 25-AUG-1999;	99US-0150566P.
PR 26-AUG-1999;	99US-0150884P.
PR 27-AUG-1999;	99US-0151065P.
PR 27-AUG-1999;	99US-0151066P.
PR 27-AUG-1999;	99US-0151080P.
PR 30-AUG-1999;	99US-0151303P.
PR 31-AUG-1999;	99US-0151438P.
PR 01-SEP-1999;	99US-0151930P.
PR 07-SEP-1999;	99US-0152363P.
PR 10-SEP-1999;	99US-0153070P.
PR 13-SEP-1999;	99US-0153758P.
PR 15-SEP-1999;	99US-0154018P.
PR 16-SEP-1999;	99US-0154039P.
PR 20-SEP-1999;	99US-0154779P.
PR 22-SEP-1999;	99US-0155139P.
PR 23-SEP-1999;	99US-0155486P.
PR 24-SEP-1999;	99US-0155659P.
PR 28-SEP-1999;	99US-0156458P.
PR 29-SEP-1999;	99US-0156596P.
PR 04-OCT-1999;	99US-0157117P.
PR 05-OCT-1999;	99US-0157753P.
PR 06-OCT-1999;	99US-0157865P.
PR 07-OCT-1999;	99US-0158029P.
PR 08-OCT-1999;	99US-0158232P.
PR 12-OCT-1999;	99US-0158369P.
PR 13-OCT-1999;	99US-0159293P.
PR 13-OCT-1999;	99US-0159295P.
PR 14-OCT-1999;	99US-0159329P.
PR 14-OCT-1999;	99US-0159330P.
PR 14-OCT-1999;	99US-0159331P.
PR 14-OCT-1999;	99US-0159631P.
PR 14-OCT-1999;	99US-0159638P.
PR 18-OCT-1999;	99US-0159844P.

Query Match 47.0%; Score 714.5; DB 3; Length 330;
Best Local Similarity 52.8%; Pred. No. 8.4e-70;
Matches 150; Conservative 39; Mismatches 82; Indels 13; Gaps 8;

QY 4	LHPPHQVAGHRASAKLGLIDGSLFYKPLQAGDRGEHEVAFYAFSAHAAVPAIR	63
Db	32 LKVPHEQVAGHIASDGKLGPLVDDQGRFFKPLQDSRGEHEAFYESFTSNMKVPDHIH-	90
QY 64	TFEPREHGTRLPTEAQPGEHPH-PLVLDDLLAGFOAPCVADIKIGAITWPPSSPREYIA	122
Db	91 RYFVYHGTQLV--EASDGGKPLHLVDDVSGYANPSVMDKIGRTWYDVSSEYFK	148
QY 123	KCLAKDRGTTSVLLGFRVSGRVVG-PEGAVMTERPEVKAMDTAGVRRVLRVYS--SV	179
Db	149 KCIKDRQTTTSLGFRVSGFKIFDHQESSFWAEKKLVLYGNADGARLALRKFSVSSNP	208
QY 180	ADEGM--DCALAAVYGGKGVLSQRELKAMFEOTLHFYFYSASILLGVDAAV-AAG	236
Db	209 ADSNLTPNCAFASVYGGCGNGLIAQLLELKDWPETQTLHFNSCSILMIYENESILMQG	268
QY 237	DGG--GVTVKLVDFAHVAEGDGVIDHNFLLGGI-CSLIKFSVDIV	277
Db	269 DDAPAPRAQVKLVDFARVLDGNGVIDHNFLLGGI-CSFIKFKDIL	312

RESULT 9
AAG24257
ID AAG24257 standard; protein; 300 AA.
XX AAG24257;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 27860.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135829P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140895P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.

PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 11-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148665P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151068P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.


```
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159337P.
PR 14-OCT-1999; 99US-0159388P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161820P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 46.1%; Score 700.5; DB 3; Length 300;
Best Local Similarity 51.8%; Pred. No. 2.6e-68;
Matches 147; Conservative 41; Mismatches 83; Indels 13; Gaps 8;

QY 4 LHPPEHQVAGHRASAKGLPLIDSGLFYKPLQAGDRGHEVAFYAFSAHAAPARID 63
Db 2 LKVPHEQVAGHTASDKGLPLVDQGRFPKPLQDSRGHEAKFYFSFTSNKKVPDHIH- 60

QY 64 TFPFPHGTRLLPTEAQGPETH-PHVLDDLLAGFOAPCVADIKIGAITWPPSSPEPYTA 122
Db 61 RYFPVYHGTLV--EASDGSGLPLHLVDVVSGYANPSVMDVKIGSRTWYPDVSEYFK 118

QY 123 KCLAKDRGTSVLGFRVSGVRVVG-PEGAVRTERPEVKAMDTAGVRRLRYVS--SV 179
Db 119 KCICEDRQXTTSLGFRVSGFKIFDQESFWRAEKKLVLYGNADGARLAKRFVSSNSP 178

QY 180 ADEGM--DCALAAVYGGKGVLSQLRELKAWFEOTLFHYFSASILLGYDAAAV-AAGG 236
Db 179 ADSNLTFCAFASEVYGGCGNIIAQLELKDWFETQTLHFNSCSILMIYENESILMKGG 238

QY 237 DGG---GVTVKLVDFAHVAGDGVIDHNFGLGLCSLIKFWSDIV 277
Db 239 DDAPAPRAQVKLVNFAHVLDDGNGVIDHNFGLGLCSFIKIKDIL 282

RESULT 10
AAG24256
ID AAG24256 standard; protein; 330 AA.
XX
AC AAG24256;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 27859.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.

XX
PR 25-FEB-2000; 2000EP-00301439.
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 14-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135623P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142053P.
PR 06-JUL-1999; 99US-0142390P.
```

```
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142902P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144351P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147193P.
PR 05-AUG-1999; 99US-0147200P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148585P.
PR 13-AUG-1999; 99US-0149368P.
PR 16-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151338P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.

PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 23-OCT-1999; 99US-0161404P.
PR 23-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161820P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 46.1%; Score 700.5; DB 3; Length 330;
Best Local Similarity 51.8%; Pred. No. 2.9e-68;
Matches 147; Conservative 41; Mismatches 83; Indels 13; Gaps 8;

QY 4 LHPPEHQVAGHRASAKLGLDGLFYKPLQAGDRGEHEVAFYFAFSAHAAVPAIRD 63
Db 32 LKVPPEHQVAGHIASDCKLGLFVDDQGRFFKPLQDGRGEHEAKFYESFTSNKKVPDHIH- 90
QY 64 TFPFRFGTLLPTEAQPGEPH-PHLVLDLLAGFQAPCVADIKIGAITWPPSSPEYIA 122
Db 91 RYFPVYHGTQLV--EASDGSGLPHLVLDVVSGYANPSVMDVKIGSRTWYPDVSEYFK 148
QY 123 KCLAKDRGTTSLVLLGFRVSGVRVVG-PEGAVMTERPERPEVKAMDTAGVRRVLRVYS--SV 179
Db 149 KCIKEDRQXTVSLGFRVSGFKIFDHQESSFWAEKKVLGYNADGARLALRKFVSSNP 208
QY 180 ADEGM--DCALAAVYGGKGVLSQIRELKAMPEQTLLFHYFASILLGVDAANV-AAG 236
Db 209 ADSNLTPNCAFASVYGGCGILAQLLEKDWPEQTQTLHYFNCSILMIYENESILMKGG 268
QY 237 DGG---GVTVKLVDFAHVAGDQVIDHNFLLGLCSLIKVSIV 277
Db 269 DDAPAPRAQVKLVNFAHVLGNGVIDHNFLLGLCSFIKFIKIL 312

RESULT 11
AAG19913
ID AAG19913 standard; protein; 286 AA.
XX AC AAG19913;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 21895.
XX
```



```
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 08-OCT-1999; 99US-0157753P.
PR 08-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159684P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161932P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 44.3%; Score 674.5; DB 3; Length 286;
Best Local Similarity 50.4%; Pred. No. 1.8e-65;
Matches 141; Conservative 42; Mismatches 96; Indels 11; Gaps 7;

QY 4 LHPPEHQVAGHRASAKGLPIIDGSLFYKPLQAGDRGEHEVAFYEAFSAHAAVPAIRD 63
Db 3 LKVPHEQVAGHIAKDGKPGPLVDKGRFFKPLQDSRGEIEVKYFESFSSNTEVPEHIH- 61
QY 64 TFFPRFHGTRLLPTEAQGEHPHLVLDLLAGQAQCVADIKIGALTWPSSPEPIAK 123
Db 62 RYFPVYHGTQAV--EGSDGA--AMVLENLARYSKPSVMDVRNGSRTPDASEEYIQK 117
QY 124 CLAKDRGTTSVLLGFRVSGVRVVG-PEGAVWRTPEVKAMDTAGVRRLRRVYS--SVA 180
Db 118 CLKKDTGTTTVSSGFRISGPEYVDHKSSFWKPKLLRGLDVGARLTLRKFVSSNLS 177
QY 181 DEGM--DCALAAVYGGKGVLSQIRELKAWFERQTLFHFYSASILLGYDAAVAGGDG 238
Db 178 DTGSKPDSAFASVYGGSHGILTQLLELKTWFENQTLYHFNCSILMYENESILKGNDD 237
QY 239 GG-VTVKLVDFAHVAEGVDHDFLGLCSLILKEVSDIV 277
Db 238 DARPOVKLVDFAHVLIDGNGVIDHNFGLGLCSFINFIREIL 277

AAG20765
ID AAG20765 standard; protein; 286 AA.
XX
AC AAG20765;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 23078.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 10-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
```

```
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142290P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144311P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144332P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 23-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145133P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 09-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 10-AUG-1999; 99US-0147935P.
PR 11-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.

PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149910P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 23-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 28-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 44.3%; Score 674.5; DB 3; Length 286;
Best Local Similarity 50.4%; Pred. No. 1.8e-65;
Matches 141; Conservative 42; Mismatches 86; Indels 11; Gaps 7;

QY 4 LHPPEHQVAGHRASAKLGPLIDGSLFYKPLQAGDRGHEVAFYEAFAHAAVPARID 63
Db 3 LKVPHEQVAGHTAKDGKPCPLVDDKGRFKFKPLQGSRGELFKVFESFSNTEVPEHIH- 61
QY 64 TFFPRFHGTRLLPTAQPGCEPHPHLVLDLLAGFQAPCVADIKIGAITWPPSPPEYIAK 123
Db 62 RYFVYHGTQAV--EGSDGA--AMVLENLLAEYSKPSVMDVMGSRRTWYDASEEYIQK 117
QY 124 CLAKDGTTSVLLGFRVSGVRVVG--PEGAVWTERPEVKAMDTAGVRVLRVYVS--SVA 180
Db 118 CLKDGTGTTTSSGFRISGFVEYDHKESFWKPKRLRLGLDVGDLRLTKRFVSSNLS 177
QY 181 DEGM--DCALAAAVYGGKGVLSQLRELKAWFEEQTLPHFYASILLGYDAAAAGGDG 238
```

```
Db      178 DTGSKPDSAFSSVYGGSHGLTQLLELKTWFENQTLHFNSCSILMVYENESILKGNDD 237
QY      239 GG-VTVKLVFAHVAEGGVIDHNFGLGCLSLIKFVSDIV 277
Db      238 DAREQVKLVFAHVLGNGVIDHNFGLGCLSFINFIREIL 277

RESULT 13
AAG06976
ID      AAG06976 standard; protein; 286 AA.
XX
AC      AAG06976;
DT      17-OCT-2000 (first entry)
DE      Arabidopsis thaliana protein fragment SEQ ID NO: 3951.
XX
KW      Protein identification; signal transduction pathway; metabolic pathway;
KW      hybridisation assay; genetic mapping; gene expression control; promoter;
XX      termination sequence.
OS      Arabidopsis thaliana.
XX
FN      EP1033405-A2.
XX
PD      06-SEP-2000.
XX
PF      25-FEB-2000; 2000EP-00301439.
XX
PR      25-FEB-1999; 99US-0121825P.
PR      05-MAR-1999; 99US-0123180P.
PR      09-MAR-1999; 99US-0123548P.
PR      23-MAR-1999; 99US-0125788P.
PR      25-MAR-1999; 99US-0126284P.
PR      29-MAR-1999; 99US-0126785P.
PR      01-APR-1999; 99US-0127462P.
PR      06-APR-1999; 99US-0128234P.
PR      08-APR-1999; 99US-0128714P.
PR      16-APR-1999; 99US-0129845P.
PR      19-APR-1999; 99US-0130077P.
PR      21-APR-1999; 99US-0130449P.
PR      23-APR-1999; 99US-0130510P.
PR      28-APR-1999; 99US-0130891P.
PR      30-APR-1999; 99US-0131449P.
PR      30-APR-1999; 99US-0132048P.
PR      04-MAY-1999; 99US-0132407P.
PR      05-MAY-1999; 99US-0132484P.
PR      06-MAY-1999; 99US-0132485P.
PR      07-MAY-1999; 99US-0132486P.
PR      11-MAY-1999; 99US-0132863P.
PR      14-MAY-1999; 99US-0134256P.
PR      14-MAY-1999; 99US-0134218P.
PR      14-MAY-1999; 99US-0134219P.
PR      14-MAY-1999; 99US-0134221P.
PR      18-MAY-1999; 99US-0134370P.
PR      19-MAY-1999; 99US-0134768P.
PR      20-MAY-1999; 99US-0134941P.
PR      21-MAY-1999; 99US-0135124P.
PR      21-MAY-1999; 99US-0135353P.
PR      24-MAY-1999; 99US-0135629P.
PR      25-MAY-1999; 99US-0136021P.
PR      27-MAY-1999; 99US-0136392P.
PR      28-MAY-1999; 99US-0136782P.
PR      01-JUN-1999; 99US-0137222P.
PR      03-JUN-1999; 99US-0137528P.
PR      04-JUN-1999; 99US-0137502P.
PR      07-JUN-1999; 99US-0137724P.
PR      08-JUN-1999; 99US-0138094P.
PR      10-JUN-1999; 99US-0138540P.
PR      14-JUN-1999; 99US-0138847P.
PR      14-JUN-1999; 99US-0139119P.
PR      16-JUN-1999; 99US-0139452P.
PR      16-JUN-1999; 99US-0139453P.
PR      17-JUN-1999; 99US-0139452P.
PR      18-JUN-1999; 99US-0139454P.
PR      18-JUN-1999; 99US-0139455P.
PR      18-JUN-1999; 99US-0139456P.
PR      18-JUN-1999; 99US-0139457P.
PR      18-JUN-1999; 99US-0139458P.
PR      18-JUN-1999; 99US-0139459P.
PR      18-JUN-1999; 99US-0139460P.
PR      18-JUN-1999; 99US-0139461P.
PR      18-JUN-1999; 99US-0139462P.
PR      18-JUN-1999; 99US-0139463P.
PR      18-JUN-1999; 99US-0139750P.
PR      18-JUN-1999; 99US-0139763P.
PR      21-JUN-1999; 99US-0139817P.
PR      22-JUN-1999; 99US-0139899P.
PR      23-JUN-1999; 99US-0140353P.
PR      23-JUN-1999; 99US-0140354P.
PR      24-JUN-1999; 99US-0140695P.
PR      28-JUN-1999; 99US-0140823P.
PR      29-JUN-1999; 99US-0140991P.
PR      30-JUN-1999; 99US-0141287P.
PR      01-JUL-1999; 99US-0141842P.
PR      02-JUL-1999; 99US-0142154P.
PR      05-JUL-1999; 99US-0142055P.
PR      06-JUL-1999; 99US-0142390P.
PR      08-JUL-1999; 99US-0142803P.
PR      09-JUL-1999; 99US-0142920P.
PR      12-JUL-1999; 99US-0142977P.
PR      13-JUL-1999; 99US-0143542P.
PR      14-JUL-1999; 99US-0143624P.
PR      15-JUL-1999; 99US-0144005P.
PR      16-JUL-1999; 99US-0144085P.
PR      16-JUL-1999; 99US-0144086P.
PR      19-JUL-1999; 99US-0144325P.
PR      19-JUL-1999; 99US-0144331P.
PR      19-JUL-1999; 99US-0144332P.
PR      19-JUL-1999; 99US-0144333P.
PR      19-JUL-1999; 99US-0144334P.
PR      19-JUL-1999; 99US-0144335P.
PR      20-JUL-1999; 99US-0144352P.
PR      20-JUL-1999; 99US-0144632P.
PR      20-JUL-1999; 99US-0144884P.
PR      21-JUL-1999; 99US-0144814P.
PR      21-JUL-1999; 99US-0145086P.
PR      21-JUL-1999; 99US-0145088P.
PR      22-JUL-1999; 99US-0145085P.
PR      22-JUL-1999; 99US-0145087P.
PR      22-JUL-1999; 99US-0145089P.
PR      22-JUL-1999; 99US-0145192P.
PR      23-JUL-1999; 99US-0145145P.
PR      23-JUL-1999; 99US-0145218P.
PR      23-JUL-1999; 99US-0145224P.
PR      26-JUL-1999; 99US-0145276P.
PR      27-JUL-1999; 99US-0145913P.
PR      27-JUL-1999; 99US-0145918P.
PR      27-JUL-1999; 99US-0145919P.
PR      28-JUL-1999; 99US-0145951P.
PR      02-AUG-1999; 99US-0146386P.
PR      02-AUG-1999; 99US-0146388P.
PR      02-AUG-1999; 99US-0146389P.
PR      03-AUG-1999; 99US-0147038P.
PR      04-AUG-1999; 99US-0147204P.
PR      04-AUG-1999; 99US-0147302P.
PR      05-AUG-1999; 99US-0147192P.
PR      05-AUG-1999; 99US-0147260P.
PR      06-AUG-1999; 99US-0147303P.
PR      06-AUG-1999; 99US-0147416P.
PR      09-AUG-1999; 99US-0147493P.
PR      09-AUG-1999; 99US-0147935P.
PR      10-AUG-1999; 99US-0148171P.
PR      11-AUG-1999; 99US-0148319P.
```


PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142772P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144864P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 44.3%; Score 674.5; DB 3; Length 327;
Best Local Similarity 50.4%; Pred. No. 2.2e-65;
Matches 141; Conservative 42; Mismatches 86; Indels 11; Gaps 7;
QY 4 LHPPEHQVAGHRASAKLGPLIDSGLFYKPLQAGDGEHEVAFYEAFAHAAPARIRD 63
Db 44 LKYPEHQVAGHTAKDGPGLVDKGRFFKPLQDGRGEIEVFYBSFSSNTEVPEHIH- 102
QY 64 TFFPRFHGTRLLPTEAQPGEPHPLVLDLLAGFQAPCVADIKIGAITWPPSPPEPYIAK 123
Db 103 RYFPVYHGTQAV--EGSDGA--AMVLENLLAEYSKPSVMDVKMGSTWYFPDASEEYIQ 158
QY 124 CLAKDRGTTSVLLGFRVSGVRVVG-PEGAVWRTERPEVKAMDAGVRRVLRVYS--SVA 180
Db 159 CLKKDTGTTTSSGFRISGFVVDHKESFVKPKERKLLRGLDVGARLTLRKVFSSNSLS 218
QY 181 DEGM--DCALAAAVYGGKGVLSQRELKAWFEQTLFHFYSASILLGYDAAAAGGOG 238
Db 219 DTGSKPDSAFASVYGGSHGILTLQLELKTWFENQTLHFNSCSILMVYENESILKGNDD 278
QY 239 GG-VTVKLVDFAHVAREGDGVIDHNFGLGSLIKFVSDIV 277
Db 279 DARPQVKLVDFAHVLDGNGVIDHNFGLGSLFINFIREIL 318

RESULT 15

ABB93390
ID ABB93390 standard; protein; 286 AA.
AC ABB93390;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 2601.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP009892.
XX
PR 28-AUG-2001; 2001WO-EP009892.
XX
PA (FARB) BAYER AG.
XX
PI Tietjen K, Weidler M;
XX
FI WPI; 2002-269010/31.
XX
DR Identifying plant target proteins for herbicidally active compounds,
XX comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms.
XX
PS Claim 5; SEQ ID NO 2601; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins (ABB90790-ABB94016)
CC for herbicidally active compounds, comprising aligning and comparing
CC nucleic acid or amino acid sequences from plant with nucleic acid or
CC amino acid sequences from non-plant organisms using suitable search
CC parameters, where plant sequences having an E-value greater by a factor
CC of 3 than the E-value of most similar non-plant sequences are selected.
CC The polypeptides or nucleic acids encoding them are useful for
CC identifying modulators. The identified modulators are useful as
XX herbicides
XX
SQ Sequence 286 AA;

Query Match 44.3%; Score 673.5; DB 5; Length 286;
Best Local Similarity 50.4%; Pred. No. 2.3e-65;
Matches 141; Conservative 42; Mismatches 86; Indels 11; Gaps 7;
QY 4 LHPPEHQVAGHRASAKLGPLIDSGLFYKPLQAGDGEHEVAFYEAFAHAAPARIRD 63
Db 3 LKYPEHQVAGHTAKDGPGLVDKGRFFKPLQDGRGEIEVFYBSFSSNTEVPEHIH- 61
QY 64 TFFPRFHGTRLLPTEAQPGEPHPLVLDLLAGFQAPCVADIKIGAITWPPSPPEPYIAK 123
Db 62 RYFPVYHGTQAV--EGSDGA--AMVLENLLAEYTKPSVMDVKMGSTWYFPDASEEYIQ 117
QY 124 CLAKDRGTTSVLLGFRVSGVRVVG-PEGAVWRTERPEVKAMDAGVRRVLRVYS--SVA 180
Db 118 CLKKDTGTTTSSGFRISGFVVDHKESFVKPKERKLLRGLDVGARLTLRKVFSSNSLS 177
QY 181 DEGM--DCALAAAVYGGKGVLSQRELKAWFEQTLFHFYSASILLGYDAAAAGGOG 238
Db 178 DTGSKPDSAFASVYGGSHGILTLQLELKTWFENQTLHFNSCSILMVYENESILKGNDD 237
QY 239 GG-VTVKLVDFAHVAREGDGVIDHNFGLGSLIKFVSDIV 277
Db 238 DARPQVKLVDFAHVLDGNGVIDHNFGLGSLFINFIREIL 277

Search completed: March 24, 2004, 20:57:40
Job time : 63 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2004, 20:56:29 ; Search time 23 seconds
(without alignments)
648.692 Million cell updates/sec

Title: US-10-042-894A-8

Perfect score: 1521

Sequence: 1 MSDLHPHQVAGHRASAK.....IKFVSIVPETHPTQLGPS 289

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	96.5	6.3	502	4	US-09-489-039A-14063
2	96	6.3	631	4	US-09-252-991A-27902
3	94.5	6.2	408	4	US-09-252-991A-32393
4	94	6.2	418	4	US-09-489-039A-14152
5	93.5	6.1	384	4	US-09-252-991A-31358
6	92.5	6.1	882	4	US-09-252-991A-17653
7	92	6.0	795	4	US-09-252-991A-19085
8	91.5	6.0	424	4	US-09-252-991A-21828
9	91.5	6.0	1175	4	US-09-252-991A-25044
10	90.5	6.0	926	4	US-09-252-991A-31053
11	90	5.9	311	4	US-09-252-991A-29142
12	90	5.9	677	4	US-09-252-991A-20406
13	90	5.9	1567	4	US-09-252-991A-20477
14	89	5.9	409	4	US-09-252-991A-31351
15	89	5.9	4551	3	US-09-320-878-1
16	89	5.9	4551	4	US-09-141-908-2
17	89	5.9	4551	4	US-09-657-440-1
18	89	5.9	4613	3	US-09-105-537-31
19	89	5.9	11877	3	US-09-105-537-6
20	88.5	5.8	370	2	US-08-846-762-81
21	88.5	5.8	474	4	US-09-489-039A-8680
22	88.5	5.8	710	4	US-09-287-849-16
23	88.5	5.8	856	4	US-09-287-849-12
24	88	5.8	463	4	US-09-252-991A-29731
25	88	5.8	485	4	US-09-252-991A-31436
26	87.5	5.8	407	2	US-08-926-327-2
27	87.5	5.8	407	3	US-09-119-918-2

ALIGNMENTS

RESULT 1

US-09-489-039A-14063

; Sequence 14063, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 14063

; LENGTH: 502

; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-14063

Query Match 6.3%; Score 96.5; DB 4; Length 502;
Best Local Similarity 21.4%; Pred. No. 0.061; 98; Indels 117; Gaps 14;
Matches 68; Conservative 35; Mismatches 98

QY	42	EHEVAFYBAFSA--HAAVPAIRIDTFFRFHGTLLPTAQGPPEPHLVDLLAGFQA	99
Db	28	ENTVVFPEAFIAQMEAFMA-----HLSEDFDFAACQR	60
QY	100	PCVADIKIGAT-----WPPSPPEPIAKLAKDR-----GTTSVLLG	137
Db	61	PIRRSIRVTLTKISVDLFLSVAPYGW-QLAPEVPCAEQFWIEREDDDALPLGSTAEHLS	119
QY	138	--FRVSGRVVGPGEAVWRTERPEVKAMDAGVRRVLRVSSVADEGDCALAAVYGG	195
Db	120	GLFIQOEASSMLPVAALFADNRQPERVMDVA-----AAPSKTQIIAARMGN	166
QY	196	KGGLV-----SQRELKA-----WFERQTLEFH-----YSASILGLGYDAVAAGDGGG	240
Db	167	AGGILANFSSASRVKVLHANSRCGISNVALTDPGRVFGAALPETFDAILLDAPCSGEG	226
QY	241	VTVK-----LVDFPA-HVABGGGVIDHNFGLGLC-----SL	269
Db	227	VVRKDADALKWSPDSNLDIAATQRELIDSFAHLPQGGTLVYS-----TCTINREENQSV	282
QY	270	IKFVSIVPETHPTQLG	287
Db	283	IEWLLSRYPQAVEILPLG	300

RESULT 2

```
US-09-252-991A-27902
; Sequence 27902, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27902
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27902

Query Match 6.3%; Score 96; DB 4; Length 631;
Best Local Similarity 23.6%; Pred. No. 0.098;
Matches 70; Conservative 30; Mismatches 82; Indels 114; Gaps 16;

QY 76 PTEAQGEPPHVLVDLLAGFOAPCVADIKIGAITWPPSSPEPYIAKC---LAKD---R 129
DB 333 PDAAAGADPGEQLV-----RCLAFVE---TDPVAPAEFGTCLRLAEQVPFV 378
QY 130 GTTSVL-----LGRVSGVRVVGEGAVWTERPEVKAMDTAGVRRVLRVYSSV 178
DB 379 GDSALVEEDPRGRFVLARHQAHHGAV-----TVPDVGAIHQRRHQVGGGLAS 432
QY 179 VA-----DEGMDCALAAVYGGKGVV-----LSQREL 206
DB 433 VAAAFDQRRRQRRRQRAVDGQGVGLAI-APVHVGVGGLRRALAGVEAIGRLTG 491
QY 207 KAWFEQTLFHFVSASILGY-----DAAVAAG-----GDGG--- 239
DB 492 LAQKEESTAAE--SRAVRLGYRGSGDGRRSVRGVAAGGQFOAGDGRGMRMGDGLLR 549
QY 240 --GVTVKLVDFAHVAE-----GDGVLDHNFVGGCLSLKFFVSDIVPEPHT 293
DB 550 CVGAAKAVGMRREARQGAAGDQGTGQVQEGHGSLSFGVRGTG---PRTPQT 602

RESULT 3
US-09-252-991A-32393
; Sequence 32393, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32393
; LENGTH: 408
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32393

Query Match 6.2%; Score 94.5; DB 4; Length 408;
Best Local Similarity 23.6%; Pred. No. 0.073;
Matches 57; Conservative 27; Mismatches 93; Indels 65; Gaps 10;

QY 53 AHAAPARIRDTFFPRFHGTRLLPTEAQGP-----HPLVLDDLLAGE----- 97

US-09-252-991A-27902
; Sequence 27902, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27902
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27902

US-09-489-039A-14152
; Sequence 14152, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709,2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 14152
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-14152

Query Match 6.2%; Score 94; DB 4; Length 418;
Best Local Similarity 25.2%; Pred. No. 0.087;
Matches 54; Conservative 20; Mismatches 70; Indels 70; Gaps 11;

QY 32 YKPLQAGDGRGEHEVAFFSAHAAVPAIRDTFFPRFHGTRLLPTEAQGPBPHPLVLD 91
DB 230 YQPLSA---EEIFAFYEVCVRHVSVPVCLYDN--PR-----TTHVMLA 267
QY 92 DLLAGFOA--PCVADIKIGAITWPPSSPEPYIAKLAQKDRGTTSVLLGFRVSGVR 144
DB 268 DELQGRIRALPAIASIKI-----PCLPAPQASE-----RVAALRQHLP 306
QY 145 --VVGPEGAVWTERPEVKAMDTAGVRRVLRVYSSVADGMDCALA--AAVYGGKGV 200
DB 307 RVTILVSGDAWA-----TAGLQAGCEAWYVCGGLFPHCSLALVRAIRSGDVA 356
QY 201 SQLRELKA--WFEQTLFHFYSASILLGYDAAA 232
DB 357 AALNEQLAPLW---RCFDRYGGSLRVITASAAA 386

RESULT 5
US-09-252-991A-31358
; Sequence 31358, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
```

```
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31358
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31358

Query Match
Best Local Similarity 6.1%; Score 93.5; DB 4; Length 384;
Matches 73; Conservative 37; Mismatches 124; Indels 107; Gaps 15;

QY 5 HPPEHQVAG-HRASASKLGLIDG-----SGLFYKPIQ----- 36
Db 82 HPGQRLGFGHALGCGPADLVHQIQLVLEIALEARAETTVVGGVEVFTLQAGKETS 141
QY 37 ----AGDRGEHEVAIFYEAFSAHAAPARIOTFFPRFHGTRLLPTEAQPGEHPHLVDD 92
Db 142 AQRVGDGDAQAT---AGGEHPGV-----LRIAGPQVRLALQADRVPVGTLEG 189
QY 93 LLAQFQAPCVADI----KIGAITWPPSPPEPIAKCLAKDRGTTSVLLGF-RVSGVRVVG 147
Db 190 FRGFAETDVADLGLHQVG-----EGADAVLGLYLRVHAVLIVQ 229
QY 148 PEGAVWRTEREVKAMDYAGVRR-----VLRVYSSVADE----GMDCLAAAVYGGKGV 199
Db 230 VQAV--GTEALQAPDGAADVRAAVDAARPGIGRVADDAELAGQEHLLAFAAQGLAEQF 287
QY 200 LSQRELKAWFEEQTLFHEFSA-----SILLGYDAAVAAGGDDGGVTVVKLV 246
Db 288 LVTVRVHVGGLIEVQPEFHGMQGGDGLLAVAGGVEVGHAAHAAEANGDGRAVAELT 347
QY 247 DFAHVAEGDGVIDHNFLGG-----LCSLIKFSVDIVETP 281
Db 348 GF-HVS-----SLAGGKENVILRSALVWIKPIITVLP 379

RESULT 6
US-09-252-991A-17653
; Sequence 17653, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17653
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (730)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-17653

Query Match
Best Local Similarity 6.1%; Score 92.5; DB 4; Length 882;
Matches 85; Conservative 29; Mismatches 99; Indels 109; Gaps 22;

QY 11 VAGHRASAKLGPL-IDSGLYFKPLQAGDRGEHEVAIFYEAFSAHAAPARIOTFFPRF 69
Db 128 VAGHQAEARVAGDLAVDAE-----QEAADR-----MPVAFAAVAI-----GTLAP-- 168

; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31358
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31358

Query Match
Best Local Similarity 21.4%; Pred. No. 0.086;
Matches 73; Conservative 37; Mismatches 124; Indels 107; Gaps 15;

QY 70 HGTRLLP-----TEAQPGE---PPHLVLDLGLAG- 97
Db 169 GAERLEPAFGLPAGLAVAGLAQAPADAVAOIAAVVEQAGERRVDFPQGLLEGGLGCG 228
QY 98 ---QAPCVADIKIGAITWPPSPPEPIYIAKCLAKDRGT-TSVLLGFRVSGV-RVVGPEGAV 152
Db 229 VFRQEQAFVDRVAGAEFRHGTSDRELVGR--LRDAGEGRAVASDIQADGVGRVV--QGAA 284
QY 153 -----WRTERPEVKAMDYA-GVRRVLRVSVSSVADEGMDCALA---AAYVGGKGVLS 201
Db 285 BEQVVRQYRVEAP-----LDAAGGVEDVV-----AAADQALDLGLAEQVRAVFGHGGDIAA 335
QY 202 Q-----LREL-----KAWFEEQTLFHEFSAIILGYDAAVAAGDGGGVTV 243
Db 336 RDARLAGDDPVGAFKRLHRVGRGFEAQV-----PGDLQAGD--GVARGHRAAGARQ 387
QY 244 KLVDFAHVAEGDGVIDHNFLGG 265
Db 388 QAADPAVAEGAAGID--FDGG 407

RESULT 7
US-09-252-991A-19085
; Sequence 19085, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19085
; LENGTH: 795
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19085

Query Match
Best Local Similarity 27.3%; Pred. No. 0.4; Length 795;
Matches 51; Conservative 16; Mismatches 70; Indels 50; Gaps 12;

QY 6 PPEHQV-----AGHRASAKLGLIDGSLFYKPLQAGDRGEHEVAFY---EAFSAHA 55
Db 188 PPVHRFRHGVRSFGRAGQG-LGILLGADGHHHR-LRRGQGHRTVHHPTDEAFRLQA 245
QY 56 AVPARIRDTFFPRFHGTRLLPTEA-----QFGEPPHVLVDLLAGFOAPCVADIKI 107
Db 246 GVRRRRRG--DREHGG---PTDAAGDGGGVHYGRDQRELRGDRQGGAD-PGSAVRL 298
QY 108 GAITWPPSPPEPIYIAKCLAKDRG---TTSVLLGFR-----VSG---VRVVGPEGA 151
Db 299 GVLDCGPRGPP-----GRSKGPAARSVERLGRCPRALVPADSVAGADRAAVALRAHA 351
QY 152 VWRTERP 158
Db 352 VLRNHRP 358

RESULT 8
US-09-252-991A-21828
; Sequence 21828, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
```


; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29142
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29142

Query Match 5.9%; Score 90; DB 4; Length 311;
Best Local Similarity 25.6%; Pred. No. 0.15;
Matches 54; Conservative 25; Mismatches 72; Indels 60; Gaps 12;
QY 7 PEHQVAGHRASAKGLPLIDGSLFYKPLQAGDRGEHEVAFYAFSAHAAPVPA-----RI 61
Db 84 PAPPAPCRAGGLALAGPAVPSRG-----AGRAGLAEVAPTQSLLAARPAVAGRRQPRL 136
QY 62 RDTFFP--RPHGTRLLPTEAQGP-----HPHL-VLDDL----- 93
Db 137 RPAYRPAGRQRRPMDRQRPGLFPQFLPQSRSLRQSPGLHAVDVVRGARRVAR 196
QY 94 LAGFQAPCVADIKIGAITWPPSPPEPIAKLAKDRGTTSVL--LGRFVSGVRVVGPEGA 151
Db 197 LADHPRPVADADHDAARRHRAEP-----DRSLAALGAAGHRLAGTEAIA--GA 244
QY 152 VWRTERPEVKAMDTAGVRRVL--RRYVSSVA 180
Db 245 VPR--RAVLEPLLEA--RRLIEPRRHGCVAA 271

RESULT 12

US-09-252-991A-20406
; Sequence 20406, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20406
; LENGTH: 677
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20406

Query Match 5.9%; Score 90; DB 4; Length 677;
Best Local Similarity 28.7%; Pred. No. 0.52;
Matches 64; Conservative 18; Mismatches 69; Indels 72; Gaps 15;
QY 5 HPPEHGVAGHRASAKGLPLIDGSLFYKPLQAGDRGEHEVAFYAFSAHAAPVPAIRD 63
Db 31 HPARHLELLRARGAIAAAPVQGRGAW-----RQGRGPDVAPGVRRAGAGH-----RHRR 80
QY 64 TFPFRPHGTRLLPTEAQGPPEPHPLVLDDLLAGFQAP-----CVADIKIGAIT--W 112
Db 81 GGVQRQH--RDLP-EAPEGAP-----ERLAGVAADPRGGQGPLC--LRGSAIDHG 127
QY 113 PPS-SPEPIYAKLAKDRGTTSVLLGFRVSGVRVVG-----EGAVWTERPEVKAMD 164
Db 128 PPRSPEPEVA-----VRLPHLLPGAAAGFGLRPRQRSGAVRR-----AVH 169
QY 165 TAGVR-----RVLRVYSSVADGMDCALAAAVYGGKG 198
Db 170 VRGVPRGDCRCDRPLRR--AAVAGAGLRPALFPVVRQRGG 210

RESULT 13

US-09-252-991A-20477
; Sequence 20477, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20477
; LENGTH: 1567
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20477

Query Match 5.9%; Score 90; DB 4; Length 1567;
Best Local Similarity 24.5%; Pred. No. 1.9;
Matches 80; Conservative 37; Mismatches 114; Indels 96; Gaps 19;
QY 9 HQVAGHRASAKGLPLIDGSG-----LFYKPLQAGDRGE-----HEVAFYEA 50
Db 1075 HEIGG---GAGVLGG--DGDGRAFOQORVLADRLQLAQGEAQEQDQADHQRHRRPADEQV 1129
QY 51 FSAHAAVPAIRD-----TFPRPHGTRLLPTEAQGPPEPHPLVLDDLLAGFQA-- 99
Db 1130 GKHRGVPSPAAGDRHGRGAAGVFLDRFHGHVPAQVDLS-GSHYP-VALGEALDEPHAT 1187
QY 100 -PCVADIKIGAITWPPSPPEPIYAKLAKDRGTTSVL---GPRVSGVRVWG---PEGA 151
Db 1188 VPALADPHRDAL-----GAERPVAVLVGDHGVAVGGIGDGRGFHQAQRRVGFARPD-- 1240
QY 152 VWRTERPEVKAMDTAG-----VRVLRVYSSVADGMDCALAAAVYGGKG--LS 201
Db 1241 -----AGEHAGGEQAFGVVQGLDHLHRLADQRVGDADARREARPRIGVDQDLE 1290
QY 202 QLRELK---AWFEQTLF-----HFYSASILLGVDAVAAGDGGGVTVKLVLD 247
Db 1291 GLAELQLGTGAFVQGEIAFCQPRIHQAGHRLGDAAVEHVAADAEELAGEGRADRL-LVD 1349
QY 248 -----FAHVAEGDGVIDHNLFGG 265
Db 1350 QRLGLAHERHVAAGDGVV-QLFLRG 1375

RESULT 14

US-09-252-991A-31351
; Sequence 31351, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31351
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31351

Query Match 5.9%; Score 89; DB 4; Length 409;

Best Local Similarity 22.9%; Pred. No. 0.3;
Matches 83; Conservative 24; Mismatches 90; Indels 166; Gaps 19;
QY 6 PPEHQVAGHR-----ASASKLGLIDSGSLFYKPLQAGD-----RGEH 43
Db 33 PPRQPARSPAVPDRSLDGOAVRLG---DGRG-----SLRGDPDRWLMLPQGLPGGTH 85
QY 44 EVAIEYAFSAHAAPAIRDTFFPRFHGTRLLPTE-AOP-----81
Db 86 -----LSGRAGLPDL-----RGHRAAPGDRAPVAATRRPENGAFSLQGEHAL 131
QY 82 ---GEPHPLV-----LDDLLAGPQAPCVADIKIGAITWP 113
Db 132 RHLGKCPESIVRPTPLGRSRRLRLVLRGDLIAQLDILLA---HPAVGRVLIGAL--- 185
QY 114 PSSPEPIAKLAKDRGTTSVLLGFRVSGV-----RVVG--PEGAVWTRTERPEVKAMDYA 166
Db 186 -----VPQVGFQLRHVAHRHVQVVTGTEAFRHRHVRVGLLDHR 227
QY 167 GVRRLVR-----YVSSVADEGMDCALAAAVYG---GKGGVL--SQLRELKAWPEEQT 214
Db 228 GGHQVILRAQBDLADLAHQVQVDVAAAAAGNDVGGQAQVIVQQLRLARQW----- 282
QY 215 LFHFYSASILLGYDAARAAAGGCGGVTVKLVDFAHVAEG-----DGVID---HNFLGGL 266
Db 283 -----VAAAHGADVAVLHQLDVAHLGIGVQRVRVDGEVEAAGGEFLGGL 325
QY 267 CSL 269
Db 326 AAL 328

RESULT 15

US-09-320-878-1
; Sequence 1, Application US/09320878A
; Patent No. 6117659
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/320,878A
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: CIP OF 09/141,908
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: 60/119,139
; EARLIER FILING DATE: 1999-02-08
; EARLIER APPLICATION NUMBER: 60/100,880
; EARLIER FILING DATE: 1998-09-22
; EARLIER APPLICATION NUMBER: 60/087,080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4551
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-320-878-1

Query Match 5.9%; Score 89; DB 3; Length 4551;
Best Local Similarity 22.1%; Pred. No. 13;
Matches 62; Conservative 31; Mismatches 106; Indels 82; Gaps 13;
QY 4 LHPPEHQVAGHRASASKLGLIDSGSLFYKPLQAGDRGEHVAIFYEAFSAHA--AVPARI 61
Db 3732 IHPALLDASLH---AIAVAGLVDBPELVRVPFH-----WSGVTVHAAGAAAARV 3777

QY 62 RDTFFPRFHGTRLLPTEAOPGEPHPLVLDDLL-LAGFQAPCVADIKIGA-----ITWPPSS 116
Db 3778 R-----LASAGTDAVLSLSLIDGEGRPLVSVVERLTLRPVTADQAAASRVGGLMHRVAMRPPYA 3833
QY 117 PEYIIAKLAKDRGTTSVLLIGFRVSGVRVVGPEGAVWTRTERPEVKAMDTAGV----- 168
Db 3834 ---LASSGEQDPHATS-----YGPTAVLGKDELKVAALLESAGVEVGLYPDL 3877
QY 169 -----RRVLRVYSSVADEGMDCALAAAVYGGKGGVLSQLRELKAWPEEQT 214
Db 3878 AALSQDVAAGAPAPRTVLAFLPAGPADGGAE-----GVRGTVAITLLELQAWLADE- 3928
QY 215 LFHFYSASILL---GYDAAAAGGCGGVTVKLVDFAHVA 252
Db 3929 --HLAGTRLLLVTRGAVRDEPGSGADDDGG-----EDLSHAA 3962

Search completed: March 24, 2004, 21:00:15
Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2004, 20:59:15 ; Search time 46 Seconds
(without alignments)

1626.915 Million cell updates/sec

Title: US-10-042-894A-8

Perfect score: 1521

Sequence: 1 MSDLHPPEHQVAGHRASAK.....IKFVSIVPETHPTQLGPS 289

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1521	100.0	289	14	US-10-042-894A-8
2	1485	97.6	289	14	US-10-042-894A-8
3	914.5	60.1	227	14	US-10-042-894A-4
4	748.5	49.2	240	14	US-10-042-894A-2
5	742.5	48.8	295	14	US-10-042-894A-14
6	722.5	47.5	310	14	US-10-042-894A-12
7	644.5	42.4	279	12	US-10-424-599-249309
8	644.5	42.4	279	14	US-10-042-894A-10
9	577	37.9	111	14	US-10-042-894A-21
10	577	37.9	111	14	US-10-094-749-2660
11	191	12.6	293	15	US-09-731-872-310
12	183	12.0	426	9	US-09-731-872-310
13	183	12.0	426	9	US-09-731-872-310
14	183	12.0	426	10	US-09-876-997-310
15	183	12.0	426	10	US-09-876-997-317

16	183	12.0	481	9	US-09-731-872-415	Sequence 415, Appl
17	183	12.0	481	10	US-09-876-997-415	Sequence 415, Appl
18	131	8.6	322	14	US-10-042-894A-23	Sequence 23, Appl
19	120	7.9	33	14	US-10-042-894A-31	Sequence 31, Appl
20	119	7.8	33	14	US-10-042-894A-30	Sequence 30, Appl
21	117	7.7	33	14	US-10-042-894A-33	Sequence 32, Appl
22	116	7.6	33	14	US-10-042-894A-32	Sequence 32, Appl
23	113	7.4	154	9	US-09-864-761-40911	Sequence 40911, A
24	108	7.1	41	14	US-10-042-894A-37	Sequence 37, Appl
25	107	7.0	41	14	US-10-042-894A-36	Sequence 36, Appl
26	105	6.9	316	14	US-10-042-894A-25	Sequence 25, Appl
27	101	6.6	756	14	US-10-156-761-11967	Sequence 11967, A
28	100.5	6.6	660	16	US-10-311-034-26	Sequence 26, Appl
29	100.5	6.6	683	12	US-10-092-900A-342	Sequence 342, Appl
30	98.5	6.5	7349	14	US-10-314-657-46	Sequence 46, Appl
31	98	6.4	41	14	US-10-042-894A-35	Sequence 35, Appl
32	98	6.4	1546	15	US-10-369-493-543	Sequence 543, Appl
33	97	6.4	41	14	US-10-042-894A-34	Sequence 34, Appl
34	93.5	6.1	756	14	US-10-080-170-424	Sequence 424, Appl
35	92.5	6.1	408	14	US-10-156-761-13474	Sequence 13474, A
36	92.5	6.1	486	14	US-10-156-761-9284	Sequence 9284, Ap
37	92	6.0	308	12	US-10-282-122A-50505	Sequence 50505, A
38	91.5	6.0	320	12	US-10-425-114-47184	Sequence 47184, A
39	91.5	6.0	281	12	US-10-425-114-46878	Sequence 46878, A
40	91.5	6.0	369	12	US-10-425-114-46333	Sequence 46333, A
41	91.5	6.0	375	12	US-10-425-114-42757	Sequence 42757, A
42	91.5	6.0	421	12	US-10-425-114-65310	Sequence 65310, A
43	91.5	6.0	425	12	US-10-425-114-57356	Sequence 57356, A
44	91.5	6.0	428	12	US-10-425-114-46479	Sequence 46479, A
45	91.5	6.0	432	12	US-10-425-114-43357	Sequence 43357, A

ALIGNMENTS

RESULT 1

US-10-042-894A-8
; Sequence 8, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1 Inositol Polyphosphate Kinase
; TITLE OF INVENTION: Genes and Uses Thereof
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; PRIOR FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Zea mays
; US-10-042-894A-8

Query Match 100.0%; Score 1521; DB 14; Length 289;
Best Local Similarity 100.0%; Pred. No. 8.2e-150;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSDLPPPHQVAGHRASAKLGPLIDGSGLFYKPLQAGDGRGHEHVAFYFAFSAHAAPAR	60
Db	1	MSDLPPPHQVAGHRASAKLGPLIDGSGLFYKPLQAGDGRGHEHVAFYFAFSAHAAPAR	60
QY	61	IRDTFFPRFHGTRLLPTEAQGPCEPHPLVLDLLAGFQAPCVADIKIGAITWPPSSPEPY	120
Db	61	IRDTFFPRFHGTRLLPTEAQGPCEPHPLVLDLLAGFQAPCVADIKIGAITWPPSSPEPY	120
QY	121	TAKCLAKDGTGTTSVLLGFRVSGVRVVGPEGAVWRTERPEVKAMDTAGVRRRLRYVSSVA	180

Db 121 IAKLAKDRGTTSVLLGFRVGVVGPAGVWMTKTERPEVKAMDTAGVRRVLRYYSSVA 180
QY 181 DEGMDCALAAAVYGGKGVLSQRLKAWFEEQTLFHFYSASILLGYDAAVAAGGDDGG 240
Db 181 DEGMDCALAAAVYGGKGVLSQRLKAWFEEQTLFHFYSASILLGYDAAVAAGGDDGG 240
QY 241 VTKLVDFAHVAEGDGVIDHNFGLGCLSLIKFVSDIVPETHPTQPLGPS 289
Db 241 VTKLVDFAHVAEGDGVIDHNFGLGCLSLIKFVSDIVPETHPTQPLGPS 289

RESULT 2

US-10-042-894A-6
; Sequence 6, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Zea mays
US-10-042-894A-6

Query Match 97.6%; Score 1485; DB 14; Length 289;
Best Local Similarity 97.9%; Pred. No. 4.6e-146;
Matches 283; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSDLHPPEHQVAGHRASASKLGPLIDGSLFYKPLQAGDRGEHEVAFYEAFAAHPAR 60
Db 1 MPDLHPPEHQVAGHRASASKLGPLIDGSLFYKPLQAGDRGEHEVAFYEAFAAHPAR 60
QY 61 IRDTFFPRFHGTRLLPTEAQPGEPHPLVLDLLAGFQAPCVADIKIGAITWPPSSPEPY 120
Db 61 IRDTFFPRFHGTRLLPTEAQPGEPHPLVLDLLAGFEAPCVADIKIGAITWPPSSPEPY 120
QY 121 IAKLAKDRGTTSVLLGFRVGVVGPAGVWMTKTERPEVKAMDTAGVRRVLRYYSSVA 180
Db 121 IAKLAKDRGTTSVLLGFRVGVVGPAGVWMTKTERPEVKAMDTAGVRRVLRYYSSVA 180
QY 181 DEGMDCALAAAVYGGKGVLSQRLKAWFEEQTLFHFYSASILLGYDAAVAAGGDDGG 240
Db 181 DEGMDCALAAAVYGGKGVLSQRLKAWFEEQTLFHFYSASILLGYDAAVAAGGDDGG 240
QY 241 VTKLVDFAHVAEGDGVIDHNFGLGCLSLIKFVSDIVPETHPTQPLGPS 289
Db 241 VTKLVDFAHVAEGDGVIDHNFGLGCLSLIKFVSDIVPETHPTQPLGPS 289

RESULT 3

US-10-042-894A-4
; Sequence 4, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase

; TITLE OF INVENTION: Genes and Uses Thereof
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Zea mays
US-10-042-894A-4

Query Match 60.1%; Score 914.5; DB 14; Length 227;
Best Local Similarity 78.5%; Pred. No. 9.4e-87;
Matches 183; Conservative 8; Mismatches 35; Indels 7; Gaps 2;
QY 1 MSDLHPPEHQVAGHRASASKLGPLIDGSLFYKPLQAGDRGEHEVAFYEAFAAHPAR 60
Db 1 MPDLHPPEHQVAGHRASASKLGPLIDGSLFYKPLQAGDRGEHEVAFYEAFAAHPAR 60
QY 61 IRDTFFPRFHGTRLLPTEAQPGEPHPLVLDLLAGFQAPCVADIKIGAITWPPSSPEPY 120
Db 61 IRDTFFPRFHGTRLLPTEAQPGEPHPLVLDLLAGFEAPCVADIKIGAITWPPSSPEPY 120
QY 121 IAKLAKDRGTTSVLLGFRVGVVGPAGVWMTKTERPEVKAMDTAGVRRVLRYYSSVA 180
Db 121 IAKLAKDRGTTSVLLGFRVGVVGPAGVWMTKTERPEVKAMDTAGVRRVLRYYSSVA 180
QY 181 DEGMDCALAAAVYGGKGVLSQRLKAWFEEQTLFHFYSASILLGYDAAVA 232
Db 180 -----CRRGDGLRARGGGVRRKRWLSIVTAAAGVVRGAASVLLGLGVDSSGL 227

RESULT 4

US-10-042-894A-2
; Sequence 2, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Zea mays
US-10-042-894A-2

Query Match 49.2%; Score 748.5; DB 14; Length 240;
Best Local Similarity 64.7%; Pred. No. 2e-69;
Matches 163; Conservative 10; Mismatches 50; Indels 29; Gaps 6;
QY 1 MSDLHPPEHQVAGHRASASKLGPLIDGSLFYKPLQAGDRGEHEVAFYEAFAAHPAR 60
Db 1 MPDLHPPEHQVAGHRASASKLGPLIDGSLFYKPLQAGDRGEHEVAFYEAFAAHPAR 60
QY 61 IRDTFFPRFHGTRLLPTEAQPGEPHPLVLDLLAGFQAPCVADIKIGAITWPPSSPEPY 120
Db 61 IRDTFFPRFHGTRLLPTEAQPGEPHPLVLDLLAGFEAPCVADIKIGAITWPPSSPEPY 120
QY 121 IAKLAKDRGTTSVLLGFRV-----SGVRVV-----GPEGAVWMTKTERPEVK 164

Db 121 IAKLADRGTTSVLLGPRVLRPRRRPRRRVADAPGCGYGHRRPPRPAALRVILP 180
Qy 165 TAG-VRRVLRVYSSVADGMDCALAAVYGGKGLVLSQLRELKAWFEQTLFHYASII 223
Db 181 TRGTARSRRCTEEKVESCHC--ASSRHGWSRLCS-----TSTRRRFFWAMM 228
Qy 224 LLGYDAAVAAG 235
Db 229 LLOSQAQEVGVG 240

RESULT 5

US-10-042-894A-14
; Sequence 14, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase
; TITLE OF INVENTION: Genes and Uses Thereof
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Parthenium argentatum
US-10-042-894A-14

Query Match 48.8%; Score 742.5; DB 14; Length 295;
Best Local Similarity 52.1%; Pred. No. 1.1e-68;
Matches 149; Conservative 45; Mismatches 79; Indels 13; Gaps 5;
Qy 4 LHPPEHGVAGHRASAKGLPLIDGSLFYKPLQAGDGEHEVAFYEAFAAHPARLRD 63
Db 2 LKAPDHQVAGHEAGLGLPLIDSGRFYKPLQDNRGSEEVAFYESSNNNIPHLR- 60
Qy 64 TFFPRFHGTRLLPTEAQPGEPPHPLVLDLLAGFQAPCVADIKIGAITWPPSSPEPIYAK 123
Db 61 KFFPIYGTIM--KASTGSDHPHVLQDLTSAHVNPSPMDIKIGSRTWAPEASEAYIAK 118
Qy 124 CLAKDRGTTSVLLGPRVSGRVVGEPA-VWTERPEVKAMDTAGVRRVLRVYSSVADE 182
Db 119 CLKXDRSTSLPLGRISGLQVYIDGSGFYKPHRNNYMRKTGPADVRLLLRKFVSNPSA 178
Qy 183 -----GMDCAAAVYGGKGLVLSQLRELKAWFEQTLFHYASIIIGYDAAVAA 234
Db 179 EMEWRTGLGPCSLASFYVGGNGHIAQIMELKTWFEDQTIYHFACSLFIPEKRLVJK 238
Qy 235 GGDGGVTVKLVDFAHVAEGDGVIDHNFGLGCSLIKFSVDIVPET 280
Db 239 GA-RSNAEVLIDFAHVDGNGVIDHNFGLGCSLIKFSIDILSET 283

RESULT 6

US-10-042-894A-12
; Sequence 12, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1el Inositol Polyphosphate Kinase

; TITLE OF INVENTION: Genes and Uses Thereof
; FILE REFERENCE: 1286
; CURRENT APPLICATION NUMBER: US/10/042,894A
; CURRENT FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,465
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Eucaalyptus grandis
US-10-042-894A-12

Query Match 47.5%; Score 722.5; DB 14; Length 310;
Best Local Similarity 52.8%; Pred. No. 1.5e-66;
Matches 149; Conservative 38; Mismatches 84; Indels 11; Gaps 6;
Qy 4 LHPPEHGVAGHRASAKGLPLIDGSLFYKPLQAGDGEHEVAFYEAFAAHPARLRD 63
Db 2 LKVPDHQVAGHRGDKGLPLVDDSGRFYKPLQSDHRGDTVEAFYFYSNTEIPGHLR- 60
Qy 64 TFFPRFHGTRLLPTEAQPGE-PPHPLVLDLLAGFQAPCVADIKIGAITWPPSSPEPIYA 122
Db 61 KFFPAFHGTXTI--EASDGGSPQPHLVLEDLVSGRTNPPLMDIKTGSRTWYPEASEEYIQ 118
Qy 123 KCLAKDRGTTSVLLGPRVSGRVV-GEPAVWTERPEVKAMDTAGVRRVLRVYSSVAD 181
Db 119 KCLEKDRNSTSVSLGRISGLRVYQNSEAGVWQPEKKVYSFNADGVRSLRKFVSNLS 178
Qy 182 EG-----MDCALAAVYGGKGLVLSQLRELKAWFEQTLFHYASIIIGYDAAVAAAGD 237
Db 179 LGPNVDPCLYASKVYCHRGGILAQLLQKWFVQVNYHYFSCSLILLYDRSALDG-- 236
Qy 238 GGGVTVKLVDFAHVAEGDGVIDHNFGLGCSLIKFSVDIVPE 279
Db 237 CAHPKVLVDFAHVMGHDGVIDHNFGLGCSVIKFIIRDADE 278

RESULT 7

US-10-424-599-249309
; Sequence 249309, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 249309
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_67156C.1.pep
US-10-424-599-249309

Query Match 42.4%; Score 644.5; DB 12; Length 279;
Best Local Similarity 50.7%; Pred. No. 1.7e-58;
Matches 142; Conservative 35; Mismatches 82; Indels 21; Gaps 9;
Qy 7 PEHGVAGHRASAKGLPLIDGSLFYKPLQAG-----DRGEHEVAFYEAFAAHPARIR 62
Db 5 PEHGVAGHKAKDGLGLPLVDDFGFYKPLQNKDDDTGSGTSLFYSLLAA-AAHDYSIR 63
Qy 63 DTFFPRFHGTRLLPTEAQPGE-PPHPLVLDLLAGFQAPCVADIKIGAITWPPSSPEPIY 121
Db 64 -SFFPAFHGTLL--DASDGGSPHPLVLEDLLCGYSKPSVMDVKIGSRTWHLGDSEDI 120

```
QY 122 AKCLAKDRGTSVLLGFRVSGVRVVGEGAVWTERPEVKAMDTAGVRRVLRVYSS---178
Db 121 CKCLKDRSSSLPLGFRISGVK--DSISSWEPTKSLQCLSAHGVALVLNKFFVSSNNI 177
QY 179 -VADEGMDCALAAAVYGGKGVLSQRLKAWFEEOITLPHFYSASILLGVDAAVAAAGD 237
Db 178 NHDDHHPDCAFATVY---GAVLERQKLDWFEVQTVVHFYSCSVLVVYEKD--LGK GK 232
QY 238 GGGVTVKLVDFAHVAEGDGVIDHNFGLGCLSLIKFVSDIV 277
Db 233 ATNPLVKLVDFAHVVDGNGVIDHNFGLGCLSFIFKLDIL 272

RESULT 8
US-10-042-894A-10
; Sequence 10, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1 Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT FILING DATE: 2002-01-09
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Glycine max
US-10-042-894A-10

Query Match 42.4%; Score 644.5; DB 14; Length 279;
Best Local Similarity 50.7%; Pred. No. 1.7e-58;
Matches 142; Conservative 35; Mismatches 82; Indels 21; Gaps 9;

QY 7 PEHQVAGHRASASKGLPLIDGSLFYKPLQAG-----DRGHEVAFYEAFAAHPARIR 62
Db 5 PEHQVAGHRASASKGLPLIDGSLFYKPLQAG-----DRGHEVAFYEAFAAHPARIR 63
QY 63 DTFPRFHGTRLLPTEAQPGCE-PPHVLVLDLQAGFQAPCVADIKIGAITWPPSSPPYI 121
Db 64 -SFPFAHGTLL--DASDGSPPHVLVLDLQAGFQAPCVADIKIGAITWPPSSPPYI 120
QY 122 AKCLAKDRGTSVLLGFRVSGVRVVGEGAVWTERPEVKAMDTAGVRRVLRVYSS---178
Db 121 CKCLKDRSSSLPLGFRISGVK--DSISSWEPTKSLQCLSAHGVALVLNKFFVSSNNI 177
QY 179 -VADEGMDCALAAAVYGGKGVLSQRLKAWFEEOITLPHFYSASILLGVDAAVAAAGD 237
Db 178 NHDDHHPDCAFATVY---GAVLERQKLDWFEVQTVVHFYSCSVLVVYEKD--LGK GK 232
QY 238 GGGVTVKLVDFAHVAEGDGVIDHNFGLGCLSLIKFVSDIV 277
Db 233 ATNPLVKLVDFAHVVDGNGVIDHNFGLGCLSFIFKLDIL 272

RESULT 9
US-10-042-894A-16
; Sequence 16, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1 Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT FILING DATE: 2002-01-09
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Zea mays
US-10-042-894A-16

Query Match 37.9%; Score 577; DB 14; Length 111;
Best Local Similarity 97.3%; Pred. No. 5.1e-52;
Matches 108; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSDLPPEHQVAGHRASASKGLPLIDGSLFYKPLQAGDRGHEVAFYEAFAAHPARIR 60
Db 1 MPDLHPPEHQVAGHRASASKGLPLIDGSLFYKPLQAGDRGHEVAFYEAFAAHPARIR 60
QY 61 IRDTFFPRFHGTRLLPTEAQPGCEPPHVLVLDLQAGFQAPCVADIKIGAIT 111
Db 61 IRDTFFPRFHGTRLLPTEAQPGCEPPHVLVLDLQAGFQAPCVADIKIGAIT 111

RESULT 10
US-10-042-894A-21
; Sequence 21, Application US/10042894A
; Publication No. US20030009011A1
; GENERAL INFORMATION:
; APPLICANT: Shi, Jinrui
; APPLICANT: Beach, Larry
; APPLICANT: Wang, Hongyu
; APPLICANT: Rafalski, Antoni J.
; APPLICANT: Cahoon, Rebecca E.
; TITLE OF INVENTION: No. US20030009011A1 Inositol Polyphosphate Kinase
; FILE REFERENCE: 1286
; CURRENT FILING DATE: 2002-01-09
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Zea mays
US-10-042-894A-21

Query Match 37.9%; Score 577; DB 14; Length 111;
Best Local Similarity 97.3%; Pred. No. 5.1e-52;
Matches 108; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSDLPPEHQVAGHRASASKGLPLIDGSLFYKPLQAGDRGHEVAFYEAFAAHPARIR 60
Db 1 MPDLHPPEHQVAGHRASASKGLPLIDGSLFYKPLQAGDRGHEVAFYEAFAAHPARIR 60
QY 61 IRDTFFPRFHGTRLLPTEAQPGCEPPHVLVLDLQAGFQAPCVADIKIGAIT 111
Db 61 IRDTFFPRFHGTRLLPTEAQPGCEPPHVLVLDLQAGFQAPCVADIKIGAIT 111

RESULT 11
US-10-094-749-2660
; Sequence 2660, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
```

APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: OTSUKI, TETSUJI
APPLICANT: WAKAMATSU, AI
APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHICO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTCYUKI
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
FILE REFERENCE: 084335/0160
CURRENT APPLICATION NUMBER: US/10/094,749
CURRENT FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/350,435
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: JP 2001-328381
PRIOR FILING DATE: 2001-09-14
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2660
LENGTH: 293
TYPE: PRT
ORGANISM: Homo sapiens
US-10-094-749-2660

Query Match 12.6%; Score 191; DB 15; Length 293;
Best Local Similarity 26.6%; Pred. No. 3.3e-11;
Matches 68; Conservative 47; Mismatches 83; Indels 58; Gaps 10;

QY 62 RDTFFP-----RFHGTRLLTTEAQGEPPHPLVDDLLAGFOAPCVADIKIGAITWPPS 115
DB 51 RKSFPWGLQCHQAHLTRLC-SEYPENKRHRFLLENVVSQYTHPCVLDLKWGTRQHGDD 109
QY 116 SPEPIYAKLADRGTTSVLLGRVSGVRVGPGEAGVWTERPEVKAMDTAGVRRVRLRY 175
DB 110 ASEKKARHMRKCAQSTACSLGVRCGMQ-----VYQTDKKYFLCKD-----KXY 154
QY 176 VSSVADEGMDCALAAVYGGK-----GGVLSQLRELKAWPEEQTLFHFYSASILLGYD 228
DB 155 GRKLSVEGPRQALYQPLFHNGSHURRELLPILLQLRALLSIIRSQSYRFYSSLLVIYD 214
QY 229 -----AAAVAAGDGGGVT---VKLVDFARVA-BG-----DGVIDH 260
DB 215 GQEPPEPAPGSPHPPEAPQAAGHSGSGGLTKVDIRMDIFAHTTYKGWNEHTTYDGP-DP 273
QY 261 NFLGLGLCSLIKFSVDI 276
DB 274 GYIFGLENLRIQLQDI 289

RESULT 12
US-09-731-872-310
Sequence 310, Application US/09731872
Patent No. US20020102604A1
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, Jean Baptiste
APPLICANT: Bougueleret, Lydie
APPLICANT: Jobert, Severin
TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
FILE REFERENCE: 78.US3.REG
CURRENT APPLICATION NUMBER: US/09/731,872
CURRENT FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 60/169,629
PRIOR FILING DATE: 1999-12-08
PRIOR APPLICATION NUMBER: US 60/187,470

PRIOR FILING DATE: 2000-03-06
NUMBER OF SEQ ID NOS: 482
SOFTWARE: Patent.pm
SEQ ID NO 310
LENGTH: 426
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: -28...-1
US-09-731-872-310

Query Match 12.0%; Score 183; DB 9; Length 426;
Best Local Similarity 25.6%; Pred. No. 3.8e-10;
Matches 62; Conservative 42; Mismatches 68; Indels 70; Gaps 7;

QY 88 LVLDLLAGFOAPCVADIKIGAITWPPSPPEPIYAKLADRGTTSVLLGFRVSGVRVVG 147
DB 204 ILLENLTSRYEVCVLDLKMGTROHGDDASEEKAANQIRKCOQSTSAVIGRVCGMQVY- 262
QY 148 PEGAVWRTPEVKAMDTAGVRRVL---RRYVSSVADEGMDCALAAVYGGK-----G 197
DB 263 -----QAGSQLMFMNKYHGRKLSMQGFKALQFFPHNGRYLRELLG 305
QY 198 GVLSQLRELKAWPEEQTLFHFYSASILLGYD-----AAAVAAG 235
DB 306 PVLKKTTELKAVLERQESYRFYSSLLVIYDGKPEVVLDSDAETLEDLSEESADESAG 365
QY 236 G-----DGGGVTVKLVDFAH-----VASGDGVIDHNFGLGLCSLIKFSVDIYP 278
DB 366 AYAYKPGICASSVDVRMIDFAHTTCRLYGEDTVVHEGQ---DAGYIFGLQSLDIVTEISE 422
QY 279 ET 280
DB 423 ES 424

RESULT 13
US-09-731-872-317
Sequence 317, Application US/09731872
Patent No. US20020102604A1
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, Jean Baptiste
APPLICANT: Bougueleret, Lydie
APPLICANT: Jobert, Severin
TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
FILE REFERENCE: 78.US3.REG
CURRENT APPLICATION NUMBER: US/09/731,872
CURRENT FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 60/169,629
PRIOR FILING DATE: 1999-12-08
PRIOR APPLICATION NUMBER: US 60/187,470
PRIOR FILING DATE: 2000-03-06
NUMBER OF SEQ ID NOS: 482
SOFTWARE: Patent.pm
SEQ ID NO 317
LENGTH: 426
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: -28...-1
US-09-731-872-317

Query Match 12.0%; Score 183; DB 9; Length 426;
Best Local Similarity 25.6%; Pred. No. 3.8e-10;
Matches 62; Conservative 42; Mismatches 68; Indels 70; Gaps 7;

QY 88 LVLDLLAGFOAPCVADIKIGAITWPPSPPEPIYAKLADRGTTSVLLGFRVSGVRVVG 147
DB 204 ILLENLTSRYEVCVLDLKMGTROHGDDASEEKAANQIRKCOQSTSAVIGRVCGMQVY- 262
QY 148 PEGAVWRTPEVKAMDTAGVRRVL---RRYVSSVADEGMDCALAAVYGGK-----G 197

```
Db 263 -----QAGSGQLMFNMKYHGRKLSVQGFKEALFQFFHNGRYLRRELLG 305
Qy 198 GVLSQLRELKAWPEEOTLFHFYSASILLGYD-----AAVAAG 235
Db 306 PVLKLTTELKAVLERQESRYFSSLLVIYDGKERPEVLDSDAEDLEDSESADESAG 365
Qy 236 G-----DGGGVTVKLVDFAH-----VAEGDGVIDHNFGLGCLSLIKFVSDIVP 278
Db 366 AYAKPIGASSVDVRMIDFAHTTCRLYGEDTVVHEGQ---DAGYIFGLQSLIDIVTEISE 422
Qy 279 ET 280
Db 423 ES 424
```

RESULT 14

```
US-09-876-997-310
; Sequence 310, Application US/09876997
; Publication No. US20030152921A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US4.CIP
; CURRENT APPLICATION NUMBER: US/09/876,997
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 310
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -28...-1
US-09-876-997-310
```

```
Query Match 12.0%; Score 183; DB 10; Length 426;
Best Local Similarity 25.6%; Pred. No. 3.8e-10;
Matches 62; Conservative 42; Mismatches 68; Indels 70; Gaps 7;

Qy 88 LVLDLLAGFQAPCVADIKIGAITWPPSPPEPIYAKLAKDRGTTSVLLGFRVSGVRVVG 147
Db 204 ILLENLTSRYEVPCLDLKMGTRQHGDDASEEKAANQIRKCCQOSTSAVIGRVCGMQVY- 262
Qy 148 PEGAVWTERPEVKAMDTAGVRVL---RRYVSSVADEGMDCAALAAVYGGK-----G 197
Db 263 -----QAGSGQLMFNMKYHGRKLSVQGFKEALFQFFHNGRYLRRELLG 305
Qy 198 GVLSQLRELKAWPEEOTLFHFYSASILLGYD-----AAVAAG 235
Db 306 PVLKLTTELKAVLERQESRYFSSLLVIYDGKERPEVLDSDAEDLEDSESADESAG 365
Qy 236 G-----DGGGVTVKLVDFAH-----VAEGDGVIDHNFGLGCLSLIKFVSDIVP 278
Db 366 AYAKPIGASSVDVRMIDFAHTTCRLYGEDTVVHEGQ---DAGYIFGLQSLIDIVTEISE 422
Qy 279 ET 280
Db 423 ES 424
```

RESULT 15

```
US-09-876-997-317
; Sequence 317, Application US/09876997
```

```
; Publication No. US20030152921A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US4.CIP
; CURRENT APPLICATION NUMBER: US/09/876,997
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 317
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -28...-1
US-09-876-997-317
```

```
Query Match 12.0%; Score 183; DB 10; Length 426;
Best Local Similarity 25.6%; Pred. No. 3.8e-10;
Matches 62; Conservative 42; Mismatches 68; Indels 70; Gaps 7;

Qy 88 LVLDLLAGFQAPCVADIKIGAITWPPSPPEPIYAKLAKDRGTTSVLLGFRVSGVRVVG 147
Db 204 ILLENLTSRYEVPCLDLKMGTRQHGDDASEEKAANQIRKCCQOSTSAVIGRVCGMQVY- 262
Qy 148 PEGAVWTERPEVKAMDTAGVRVL---RRYVSSVADEGMDCAALAAVYGGK-----G 197
Db 263 -----QAGSGQLMFNMKYHGRKLSVQGFKEALFQFFHNGRYLRRELLG 305
Qy 198 GVLSQLRELKAWPEEOTLFHFYSASILLGYD-----AAVAAG 235
Db 306 PVLKLTTELKAVLERQESRYFSSLLVIYDGKERPEVLDSDAEDLEDSESADESAG 365
Qy 236 G-----DGGGVTVKLVDFAH-----VAEGDGVIDHNFGLGCLSLIKFVSDIVP 278
Db 366 AYAKPIGASSVDVRMIDFAHTTCRLYGEDTVVHEGQ---DAGYIFGLQSLIDIVTEISE 422
Qy 279 ET 280
Db 423 ES 424
```

Search completed: March 24, 2004, 21:04:54
Job time : 47 secs